VOLUME 22 DECEMBER, 1958 NUMBER 4

Progress in Parapsychology.....*Gardner Murphy* 229

Scaled Measurement in Psi Research.....*R. A. McConnell* 237

A Survey of Work on ESP and Teacher-Pupil Attitudes.....*Margaret Anderson and Rhea White* 246

Preliminary Experiments on Psi Phenomena with Relatively Unconscious Participation
Ronald E. Shor, Laurence S. Youell, and Hershel Berkowitz 269

Continuous Variable Trials.....*R. A. McConnell* 279

Book Reviews

Telepathie en Helderziendheid (Telepathy and Clairvoyance) by W. H. C. Tenhaeff.*J. J. Poortman* 291

Psi Cognition by K. Ramakrishna Rao
Stephen I. Abrams 292

Parapsychological Abstracts..... 294

Announcements by the Parapsychological Association..... 297

Glossary 306

ENTIFIC QUARTERLY DEALING WITH EXTRASENSORY PERCEPTION, THE PSYCHOKINETIC EFFECT, AND RELATED TOPICS

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

EDITORIAL STAFF

J. G. PRATT, L. E. RHINE, and W. G. ROLL, *Editors*
DOROTHY H. POPE, *Managing Editor*
J. A. GREENWOOD and T. N. E. GREVILLE, *Statistical Editors*

This JOURNAL is published on the thirtieth day of March, June, September, and December by Duke University Press, Durham, North Carolina.

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, North Carolina. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, the JOURNAL cannot guarantee that they will not be lost in transit, and contributors are urged to keep copies of their papers. All contributions should be typewritten, double spaced. References should be given in the form adopted by the JOURNAL.

Reprints may be ordered when the proof is returned. The JOURNAL will bear one half the cost of reprints up to two hundred copies. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to Dorothy H. Pope, College Station, Durham, N. C.

The subscription price is $6.00 a year; single current numbers $1.50. The rate for back volumes is $7.00; for single numbers $1.75. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. All remittances should be made payable to the JOURNAL OF PARAPSYCHOLOGY.

Subscriptions may be obtained in Great Britain through the Cambridge University Press, Bentley House, Euston Road, London, N. W. 1.

Entered as second class matter May 21, 1942, at the post office at Durham, N. C., under the act of March 3, 1879.

DUKE UNIVERSITY PRESS

© Journal of Parapsychology 1958

The Journal *of*
Parapsychology

VOLUME 22 DECEMBER, 1958 NUMBER 4

PROGRESS IN PARAPSYCHOLOGY[1]

By GARDNER MURPHY

IT IS GOOD TO SEE so many faces of old and new friends and collaborators, each constituting something vital: Jan Ehrenwald with the urgently needed depth approach; pioneer experimenters with new methods like George Estabrooks and Gaither Pratt; vigorous young beginners like Stephen Abrams; my treasured research companion from the Harvard of 1942, Gertrude Schmeidler.

I

We are gathered to advance the study of parapsychology. I am grateful indeed that you have honored me with an opportunity to address you on this problem. The only task that seems to me worthy of the occasion is that of asking how we may more wisely, rapidly, effectively improve our research. I say this partly because I feel a serious risk of stagnation, preoccupation with minor problems and a dearth of big ideas in our field. I admit, in relatively honest moments, that my own lack of incisive imagination and an increasing inability to create new perspectives has irked me and made me feel guilty, and in the same breath that I believe this unimaginativeness is characteristic, more or less, of our whole movement. Philosophical breadth is important; but it is not the same as research creativeness. I shall attempt to throw myself bodily at the target of research creativeness or lift myself by those bootstraps which will take me out of the sluggishness of today;

[1] Paper read at the Convention Dinner of the Parapsychological Association on September 5, 1958.

and if I fail I shall at least lunge toward a target which it will be your task to define more clearly.

II

I have six specific suggestions to offer. (1) First, that we study more closely the history of science, and note the peculiar attitudes of the great movements in which science breaks with its past, striving to see how the evolution of parapsychology might benefit from awareness of the nature of those great turning points. (2) Second, I shall ask about the role of scientific models, and ask whether anything is wrong with our models in parapsychology. (3) Third, I shall ask you to think with me about models which utilize our knowledge of biology. (4) Fourth, I shall try to raise the same question about the social-cultural context of our work, conceived in terms of a field-theory which provides the context within which paranormal processes may occur. (5) Fifth, this will lead to the question whether hypothesis-construction and model-building can be undertaken and suitably tested within that larger framework which comprises both the biological and the cultural context of paranormal events, and ask whether the problem of confirmability or repeatability, the genuine experimental control of paranormal processes, is ready for formulation in these same terms. (6) Sixth, I shall speculate on the question whether we are consistently applying what we know as psychologists. Specifically, I shall ask what troubles may be arising from conceptions which may no longer be appropriate in psychology, and I shall stress our need to use models which are appropriate for the half-formed, half-determined, probabilistic nature of cognitive, affective, and conative functions in response to a half-formed world, which throw upon the living organism the need to use whatever paranormal functions it possesses.

III

1. To begin then with the history of science. A moving object continues along a straight line, as Newton showed, until deflected. The same is true of an intellectual movement. To predict, you must extrapolate. Science does more and more of what it is

already doing. But ultimately another trend, dealing with the phenomena in a different way—for example, optically rather than mechanically—interacts with the first; and there is an *emergent,* a new thing in the world. Examples from psychology are the interaction of Darwinism with experimental physiology, as in the work of Walter Cannon, and of psychoanalysis with physiological optics, as in the movement known as the "New Look," appearing first, perhaps, in Hermann Rorschach's ink blot test and in Harry Murray's Thematic Apperception Test. Are we using this principle of emergence, or just coasting in straight lines? As I read the Journals, I see dozens of attempts to proceed by extrapolation from the experiments of Rhine; but almost no cross-fertilization by combining Rhine, let us say, with Tyrrell or Whately Carington, or Freud, or James. What would happen if somebody read James on Mrs. Piper's trance and saw its relation to experimental precognition?

2. Secondly, almost everybody today is convinced that psychology has done poorly to go on with the hypothesis-construction and hypothesis-testing that were appropriate a half-century ago. Almost never does a single independent variable stand forth, neatly permitting everything else to be held constant while we measure the dependent variables. We deal, ordinarily, with a matter of interdependent components in a complex situation, and we must use the Fisher methods, especially when we deal with interactions. But have we in parapsychology learned this lesson? While biology and psychology have had to admit that they cannot continue to go on with the single-track approach, most of our work in parapsychology is still one-variable work. We need perspective and practice in model-building and model-testing.

3. Third, I cannot personally entertain much hope of major breakthroughs without an attempt at definition of the complex multi-dimensional biological contexts in which paranormal events occur. This may be thought premature. But we are already bogged down by the large number of untestable hypotheses and the fact that those few which pay off lead, as a rule, to no experiment confirmable or repeatable by others. Plainly, complexity of models is demanded by complexity of events in science. And there is no

way to discover deficiency in models that is half as good as to construct them in detail and see if they are at least consistent and oriented to such reality as we know. Take a simple biological conception, very good so far as it goes, that certain drugs facilitate ESP. But one of the first things in all drug research is the richness of individual differences. Take a beautiful study like Roger Williams' book on *Biochemical Individuality* and you note that every biochemical event is related to an individual biochemical system. Studies of drugs, toxins, foods, oxidation, remind us that a suitable model for biochemical research today must be based on the best view of the individual living system that we can obtain. The same is true for every biological component in our study of our subjects.

4. The same conception of multiple interacting factors forces itself upon us with a vengeance when we deal with the sociocultural contexts of paranormal events. There is still almost no cross-cultural research in parapsychology, and even within our own culture, the studies of interpersonal factors in the parapsychological laboratory are limited to a few studies of subject-experimenter relationships. But the relationships are certainly dependent upon the family and community background of subject and of experimenter. Of course, if the results we get are consistent, we can for the time being emphasize what we can control, but if they are inconsistent, they presumably arise partly from sociocultural contexts that have not yet been analyzed.

5. As is already happening in general psychology, such an analysis will certainly force us into the development of models in which biological and cultural components are viewed in their interaction. If we can attempt the same in parapsychology, it will profoundly influence our way of thinking about our research methods. I suspect indeed that our deficiency in inventing new methods (for there have been almost no new methods at all since those invented by Rhine twenty or more years ago) may be due to deficiency in models. It is only when baffling conceptual problems loom upon us that we have the courage and the imagination to invent a new method. Yet paradoxically much of the history of science is a simple matter of finding new methods. What happened to astronomy when Newton and Leibnitz invented the differential and integral calculus;

to biology when Leeuwenhoek invented the microscope, to medicine when the Hertzian waves became a method of examining the interior of the body? Yes, something vigorous must be done to launch us on a bender of method-inventing. I know of nothing that can do it but the discovery of the deficiency of our models, as shown in the non-confirmability of so many of our reports.

But why is it, in the meantime, that we do not see with open eyes how inadequate our models are? I think the answer lies mainly in our manner of looking at this problem of confirmability. The fundamental rule in laboratory science is that you truly have captured a phenomenon and begin to understand it only when you can so fully specify the conditions which engender it, that you can yourself make it happen again and again, and other qualified workers can do the same. We do not, for the most part, even attempt this in parapsychology. The few cases such as the Schmeidler, van Busschbach, and Anderson-White studies, that seem promising, are indeed important but they are very far from achieving that standard of confirmability which we expect when we hear the word science, and it is clear from these cases that there are at least a half-dozen variables, personal and interpersonal, that will have to be understood and controlled before we can write a consistent statement of what happens. In the meantime there are literally dozens of new one-variable studies which, even if they succeed, can serve only to dilute the urge to create a science, a discipline, a coherent system of ideas out of parapsychology.

6. But perhaps our persistence in making certain methodological assumptions, to which we cling despite cumulative evidence that they are errors, is due to our habit of not using the psychology that we know. One assumption relates to the supposed basic uniformity of the effects of experimenter attitudes; another, the supposed uniformity of the effects of subject attitudes, particularly college student attitudes. Assumptions about the uniformity of the attitudes of subjects in experimental situations don't work well even in simple tasks like the learning of nonsense syllables, and certainly work very badly in the delicate interpersonal situations involved in paranormal interaction. If psychiatrists are learning that patients' responses in psychiatric hospitals are expressions of very complex field relation-

ships, perhaps subjects in parapsychology experiments will ultimately be understood to require at least this much sophistication. We are dealing with plastic cumulative material in which time trends as well as cross sections have to be kept constantly in mind. Even in the more sophisticated studies the deep-level needs and attitudes of experimenters and subjects are but dimly understood, and of course yield no consistent predictable results. A combination of the anthropologists' and the clinical psychologists' skill might help us to understand who those people really are whom we scoop up in our experimental net, and might deal then with full-bodied people.

Another assumption which we would make less frequently if we functioned as psychologists is the assumption that every paranormal cognitive process is inevitably a form of perception rather than a form of recall, recognition, phantasy, conceptual thought, or what not. Even the use of the term paranormal cognition, awkward as it is, might remind us that we are constantly dealing with more than perception. A humble but important illustration of this came up in connection with Charles Stuart's reminder that two things can be paranormally related because they are alike or because they are associated, rather than because both constitute an expression of the same event. If you are trying to transmit to me the word *hat,* my response *head* may, under certain specific conditions, be the expected paranormal response; or to the word *tennis* the appropriate response may be *golf* or *ball* or *love* or *set,* depending upon a variety of contexts ordered around the word *tennis.* We know this in the study of aphasia. Are we sufficiently aware of the principles of general psychology to recognize that a paranormal process may be very far from a carbon copy of the material indicated in the target? This principle was, of course, stated at the Duke University laboratory many years ago, and was a cardinal feature of one phase of the work of the late Whately Carington, but I look in vain as I read the Journals for evidence that this way of thinking has gotten into our blood. I fear in fact that the use of the term psi, useful as it is in some connections, is making us forget the difference between extrasensory perception and other paranormal cognitive functions.

Another issue that makes me wonder whether we are really working as psychologists is our tendency to set up for our subject

tasks which are essentially mechanical in nature, in which there is no initiative, no creative exploration. This can at times become a violation of the reality of the paranormal process. The paranormal process is evidently a weak biological function, making fleeting evanescent contact with reality in its own way, almost never like the sharp, clear, exteroceptive function of seeing and hearing. The paranormal processes may apparently in some cases be standins or substitutes for exteroceptive function, and in some cases very possibly represent "potentiations" in which the exteroceptive function at a low level is eked out and abetted by the paranormal process. In order to set up what we consider a proper experiment, we ignore this possibility and assume that regardless of the situation, we must completely exclude sensory cues. Now for purposes of certain kinds of elementary statistical analysis, the exclusion of sensory cues is a fundamental and essential first step. But if there should be very rich interactions, in R. A. Fisher's sense, between ordinary sense perception and extrasensory functions of various sorts, we should mutilate or exclude these altogether by our usual methods. We have talked now for many years about the potentiation in which a weak paranormal process is grafted upon a weak normal perceptual process, in which we compare the perceptual process working without ESP with what occurs when the ESP factor is added. Yet, as far as I know, two hesitating and tentative little studies are the only ones in which this principle is being expressly formulated.

Still another example of our lack of contact with psychology: One psychological principle which has spread through the psychological world like wildfire in the last ten or a dozen years is the conception of game theory, emphasizing the hazard element or aleatory element involved in all transactions with the environment. This idea is related of course to information theory on the one hand and to Brunswik's conception of probabilistic functionalism on the other hand. The organism wanders through this tangle or undergrowth which we call the world of the senses, piecing together fragments of usually unrepeatable evidence in endlessly varying patterns, and though bungling a good deal, still making some sort of contact with reality. One of the basic functions of the organism is to shift the P-values constantly, relying sometimes on evidence

in which the P-value is one-half, that is, there is an equal likelihood of a door being opened or closed, and in other cases where P-values are of the order of one in a thousand, where for example a cry for help or the sound of a siren is just one out of literally a thousand different stimuli that might hit us, so that the significance for the individual of the siren or the cry for help is enormously greater than it could be if P-value were one-half. The constant utilization of different P-values in empirical attunement with cultural and individual idiosyncracies is one of the major things which should preoccupy the parapsychologist. One reason for this lies in the basic trueness to life of the shifting P-value model. Another equally important reason is that for the most part, with the faint extrasensory powers of today, a large proportion of our experiments reach only marginal significance and on repetition do not again achieve this marginal significance. Perhaps the use of a lower P-value would alert us to those relatively rare but important things which indicate a true function, standing out crisply against a background and moving with a swifter replication. First impressions could later yield the solid outlines of a true paranormal contact with reality. Empirical manipulation of P-values by using different kinds of target materials is one of the most fundamental ways of getting into parapsychology the true spirit of our openness in relation to a vague and challenging environment. This happens to be a useful statistical principle; but more important, it is a useful psychological principle.

IV

So you see, I am pleading here for more self-criticism, more sophistication, more biological-mindedness, more cultural-mindedness, more psychological-mindedness. The fact that there are only a few of us, with no great funds coming our way, is indeed too bad. But the initial work of Rhine, Carington, and Soal did not call for large numbers, nor for much money. What was called for was ideas. To the younger people especially I would say: older people can do a little to help you, but today and in the future you are the fabricators and the processors of ideas; it is your ideas that will catalyze the growth of parapsychology. It is your turn now.

Menninger Foundation
Topeka, Kansas

SCALED MEASUREMENT IN PSI RESEARCH[1]

By R. A. McConnell

THE BUSINESS OF SCIENCE is the understanding of relationships among observables. Scientific relationships may be quantitatively precise, as between the position and time of a falling object, or descriptive and highly abstract, as the evolutionary connection between two geologic fossils.

The branches of science which have had great cultural impact are for the most part those in which relationships have been successfully quantified. There is even a tendency to insist that purely qualitative study is not entirely scientific. This is epitomized by a famous statement by the 19th century physicist, Lord Kelvin. He said: "When you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of *science,* whatever the matter may be."

I would like to examine the applicability of this idea to parapsychology. If I present a physicist's point of view, it is also a viewpoint now largely adopted by psychology. The dictum, that observation should, where possible, be measured on a scale, is not likely to be abandoned, even though many other conceptions borrowed from physical science may ultimately prove ill-adapted to the study of biological materials.

The idea of graduated measurement is inextricably bound up with recent progress in all of the life sciences; and if in psychology, or in any part of it, we wish to forego this idea at least as an ideal toward which we aim, we must first conceive and prove-out a substitute methodology. This would be a revolutionary undertaking for

[1] Presidential address presented at the meeting of the Parapsychological Association on September 5, 1958.

This paper is indexed as Publication No. 58 in the Biophysics Department of the University of Pittsburgh.

which I see among our membership neither conscious willingness nor technical preparation.

Indeed, the desirability of scaled measurement is so obvious to all of us that I must begin with an apology. None of the ideas I shall express is new; although a full appreciation of their applicability is prevented by the experimental difficulties that plainly stand in our way.

The simplest organizing scheme for the explorer in a new scientific field is to categorize observations. Differences may be recognized by the use of different conceptual boxes.

Sometimes the boxes are related to one another obscurely if at all; as for example, when one separates flowers by their color, or spontaneous psi experiences by content.

Sometimes an explorer's categories can be arranged in the order of increasing amounts of the same quantity, as for example, the size of flowers, or the number of hits in an ESP run. Whenever ranking by magnitude is possible, one tries to abandon simple categorization and to substitute a measuring scale.

In order to develop a measuring scale for psi phenomena, how might we proceed? Most past experimental work has been done with multiple-choice tasks (such as the guessing of one out of five possible ESP symbols) or by tasks scored by multiple-choice methods (such as the matching of ESP drawings against target pictures). To develop a scale we can either aggregate similar multiple-choice scores until an average is obtained, or we can alter the experimental task into something intrinsically continuous, such as the dial reading on an old-fashioned voltmeter.

I shall say more about continuity, but at this point I want to stress the idea that most scaled measurements imply continuity, at least over a limited range, and at least in conception if not in observation. For example, if we say that a subject had an average ESP score of 5.4 in an experiment with the standard ESP cards, we must mean that we would like to think of an ESP card-guessing ability which is essentially continuous in nature. On the scale of such an ability a score of 5.41 would be a little higher than 5.40, admitting, of course, that statistical uncertainty would limit the precision of observation in any particular experiment.

The conception of continuity may exist even where steps on the measuring scale are unavoidably fixed. This might occur, for example, when scoring a psychological test by a merit-point system in which the maximum number of points is pre-determined. With such a scale, the obtained score represents only some unspecified combination of item-successes. That is, a score of three-out-of-ten does not tell which three, but only that some three items were correct. In any theoretical interpretation of a score of this kind, one is forced to use the concept of continuity because the steps themselves are meaningless, except in a technical or tautological sense. Now scores of this seemingly discontinuous, but actually continuous, kind are very common although I suspect that psychologists are not always aware of the implications of functional continuity.

I will not deny that there may be psychological measures which are scaled in the sense of representing progressively greater or lesser quantities, but which are discontinuous in conception. However, none such has occurred to me, and I dare say that they are a rarity.

To avoid misunderstanding I shall discuss briefly two ways in which incidental discontinuities may appear while describing a continuous quantity. By incidental discontinuities, I mean those which have no psychological significance even though they may be of technical importance.

It is well known that the use of a stepwise scale for the representation of a continuous concept causes grouping errors in statistical analysis. We shall assume that these errors are recognized and suitably corrected and are of no special interest in all of the discussion which follows.

And likewise, if the conceptually continuous variable is the limiting case of the average score for an aggregation of discontinuous trials, only certain values of the variable will be observed in a real experiment. For example, if one gives exactly ten ESP runs to each subject, the observed run averages will always appear in exact tenths with no scores in between. Again, the implications of this fact are obvious and need not distract us from a methodological discussion of continuity.

Except as the average of a number of discontinuous binomial trials, continuous variables are few in the literature of parapsychology. The clock card introduced by Fisk and Mitchell (1) is the only continuous ESP variable that has been widely used. Instead of the standard five-symbol ESP deck, these experimenters used a clock face with a single hand pointing to one of the 12 clock hours. The subject was asked to guess a preselected hour appearing on the card, and he received a score that was linearly related to his time error.

Although the clock-card response error is limited to 12 possible values, it approximates a continuous variable well enough for most purposes. The clock cards can also be scored discontinuously. That is, a guess can be treated as a success or failure with a binomial probability of one-twelfth. This raises an interesting question. With the same data, which scoring is more sensitive to ESP, the continuous or the discontinuous? I have discussed this question elsewhere. (2)

For the development of scientific understanding, the observation of a single variable is not enough; we must concern ourselves with the relationship between two or more. We would like to know how the operation of psi depends upon various factors, such as the personality of the subject or his physiology, upon the kind of psi-testing apparatus that we use, upon the structure of the data record, or upon any of a vast number of environmental conditions. Most of the variables we might investigate will prove to be of no importance, and rough methods of observation may suffice to tell us so. But where an independent variable is believed to affect the emergence of psi, we should begin to think in terms of scaled measurement.

Much of the advantage of scaled measurement is lost unless there are at least two scaled variables in the experiment. One cannot determine a complex functional relationship if one variable is continuous while the other is merely present or absent. Let me illustrate this by an example.

The first successful scaled PK experiments—at least in modern times—are those by Forwald (4). He measures the distance a die is deflected to the right or left in rolling to rest. Thus, he

substitutes a measurement of linear distance for a choice among six sides of the die. But the second variable, the wishing of the experimenter, has not been quantified: he simply wishes the dice to go right or left.

Methodologically, the next step—and a most important one—will be to search for continuous variables which affect the die displacement. Forwald believes he may have found such variables in the thickness and atomic number of the coverings he places over his dice. If the relationships shown in his still unpublished data prove real rather than psychological artifacts, he will have made a most important contribution.

To the best of my knowledge there is only one kind of functional relationship between continuous variables which has been clearly demonstrated in parapsychology—and even that is of a limited nature. I am thinking of position effects in the data record, and particularly of the decline effect. In some experiments, where the position effects are strong and the data arrangement suitable, one can trace the variation in average score through successive segments of the run or data page and demonstrate a continuous trend. Obviously, it takes at least three adjacent data positions before we can think in terms of even a linear continuous function, and five to ten positions before we can have much confidence in our thinking.

What is the scientific value of scaling a functional relationship between two variables? To answer this question, let us assume that the relationship is a line of some sort, and that it is known with enough precision so that we can perceive its continuity and curvature.

Under these conditions the continuity of the relationship gives assurance of validity, beyond that offered by simple repetition of observation. In a scientific experiment it is often difficult to know whether we are changing extraneous, unknown variables along with the x and the y in which we are interested.

But if a little change in x causes a corresponding little change in y, and if similar pairs of incremental changes occur over a range of x and y values, we can have more confidence in the interdependence of x and y than we could, say, in another experiment where

x and y could be either present or absent but had no measurable magnitude.

We can express this idea in the language of information theory by saying that knowledge of a continuous function implies more "information" than does knowledge of a simple conditionality function. And with more information, goes more confidence.

With more information, there also goes more understanding. A continuous functional relationship allows us to unravel complex situations, and by now we suspect that parapsychology is full of complexities.

It is unfortunately true that extreme complexity between two variables appears as discontinuity, whether it is or not. If a relationship is complex, its continuity or discontinuity will not be evident unless the values of the independent variable are sufficiently close-spaced. In other words, a jagged line based upon a scattering of points may be difficult to interpret, but a smooth curve with only isolated discontinuities can be scientifically revealing.

An interesting question which may arise is whether complex functional relationships can exist, or, if found in data, whether they are real. Under what circumstances, for example, should one expect, or accept, a pronounced discontinuity in a measured relationship between ESP scoring and some independent variable?

There is no doubt that many physical measurements are discontinuous. A wooden plank will bend slowly until suddenly it breaks. Likewise with human behavior. If you persistently annoy your bridge partner, he may smile up to the very moment that he kicks you in the shin. It is not safe to say in advance whether he will lose his appearance of good humor slowly or all at once. Nor can we make a better prediction in the case of psi research. Any kind of complex relationship ***might*** exist between successful card guessing and other variables, even when a simple relationship would seem more reasonable to most experimenters.

But to say that a complex relationship may exist is not to say that we can observe it. I want to examine with you some of the limitations that we should impose upon our expectations when we set out to discover new functional relations in parapsychology. If there are no limitations upon what is psychologically possible, there

may be some upon what is mathematically possible with the kind of measuring tools we are accustomed to use.

Up to now I have discussed a functional relation as a given curve without regard to uncertainties in the two variables. Yet most of the variables we may think of will have rather large inherent errors of measurement. ESP and PK scores usually have standard deviations close to the theoretical chance values. That is to say, the uncertainty in the measurement of ESP ability in a test with several hundred card trials is usually many times the causal change to be expected from a small change in any independent variable. Nevertheless, if the change in the independent variable can be measured precisely, we can often continue to gather data until the standard deviation in score average is so small that the functional change is shown to be meaningful.

Let me give an example from an analysis for position effects. Consider as an independent variable the 25 positions in the standard run of cards. Each of the positions can be measured with absolute accuracy; that is, the first card of a run is always the first card and can never be confused with the second. Hence, it is legitimate to average all the first guesses and compare them with the second, and if the psychological conditions can be maintained long enough, a statistically significant scoring difference should be discovered between the first and second trial positions in the run.

Now let us think of another example that leads to a different conclusion. In their paper, "The Exploration of ESP and Human Personality" (3), Nicol and Humphrey measured ESP performance as a function of certain personality traits using a number of subjects. The personality traits were measured by a scaled score derived from the answers to a questionnaire. Hence, quite properly, these experimenters looked for a functional relationship between ESP performance and each personality trait.

Now the difference between this and my previous example is that this time the independent variable is not precisely known. The personality trait score has its own statistical uncertainty, and this cannot be reduced by repeated measurements on the same subject.

Consequently, it is not entirely correct to average together all

ESP scores which have a particular personality score. The mean ESP score which one obtains in so doing corresponds to a range of personality scores. Although the blurring in the y-variable is reduced, the blurring in the x-variable remains.

(I hasten to point out that Nicol and Humphrey did not average the ESP scores from different subjects. The scattering of their individual points was such that they attempted only a linear correlation between variables.)

These examples point to two extreme kinds of situations. If the independent variable can be measured with perfect accuracy, we can discover complex relationships merely by gathering enough data, regardless of the uncertainty in the dependent variable. But if the independent variable can be determined only once for a given test entity (the subject) and has a standard deviation of measurement roughly equal to or larger than 20 per cent of its total range, no amount of averaging will justify fitting anything but a straight line to the data.

Where the independent variable is an ordinary psychological measure, the situation is likely to be somewhere between these two extremes. For example, it is not at all clear in the case of Nicol and Humphrey's work that the standard error of their personality trait measurements was so large as to have precluded a second-degree curve, had they gathered more data than they did.

What useful conclusion might we draw from these considerations? First, the amount of information about the dependence of psi upon other variables will be limited by the uncertainties in those other variables as well as by the variation in ESP ability. If we hope to discover curved or discontinuous relationships with ESP, we should choose independent variables that can be measured with sufficient precision.

Secondly, if we gather many data and find statistically significant discontinuities or fluctuations in averaged y-values, we should test those fluctuations against the standard deviation of the x-variable to see whether they are believable or whether they are probably artifacts.

As we look to the future, we will attempt more and more to use the scaled measurement of variables in parapsychological re-

search. We should not underestimate the difficulty of this task. It may be that psi phenomena will not lend themselves to quantification. This, however, is a question still to be answered; up to now we have scarcely tried.

References

1. Fisk, G. W., and Mitchell, A.M.J. ESP experiments with clock cards: A new technique with differential scoring. *J. Soc. psych. Res.* (London), 1953, **37**, 1-14.
2. McConnell, R. A. Continuous variable trials. *J. Parapsychol.*, 1958, **22**, 279-90.
3. Nicol, J. F., and Humphrey, B.M. The exploration of ESP and human personality. *J. Amer. Soc. psych. Res.*, 1953, **47**, 133-78.
4. Pratt, J. G., and Forwald, H. Confirmation of the PK placement effect. *J. Parapsychol.*, 1958, **22**, 1-19.

Department of Biophysics
University of Pittsburgh
Pittsburgh, Pennsylvania

A SURVEY OF WORK ON ESP AND TEACHER-PUPIL ATTITUDES

By MARGARET ANDERSON AND RHEA WHITE[1]

ABSTRACT: This report concerns nine exploratory studies which were carried out in 1957-58 on the relationship of teacher-pupil attitude and level of ESP scoring. In view of the experimental variations from the original A-W design, the authors are presenting these studies in abbreviated form. Though some of the results were similar to earlier findings, the scoring differences did not reach the level of statistical significance. The authors also give a complete summary of their work to date on the relationship of ESP scoring and teacher-pupil attitudes. They present this full account of their work preparatory to a forthcoming publication on the relationship of ESP and student class grades which is based in part upon it.—Ed.

INTRODUCTION

THE EXPERIMENTS of van Busschbach (**7**, **8**, **9**) as well as our own (**1**, **2**, **3**, **4**, **5**) have produced evidence of ESP in the classroom. Ours have also indicated some of the conditions under which positive or negative deviations from chance expectancy might be anticipated.

Our original school experiment (A-W 1) was conducted in seven academic classes at the secondary school level in seven different states. The experiment was then successfully carried out with two fifth and two sixth grades in an elementary school in the Midwest. This, in turn, was followed by a replication of our first research at the secondary school level.

In an effort to get a broader perspective of the demonstrability of psi in the classroom, we have carried out 11 exploratory experiments divergent in design from our original research. We believe that parts of each of these studies might bring the experimental

[1] The authors wish to thank Mrs. Mary Davies, Principal of the Hawthorne Elementary School in Wheaton, Illinois; Mrs. Merta Reed of St. Charles, Michigan; Miss Sally Springer of Moberly, Missouri; Mr. Claude Burns of San Diego, California; Superintendent Miles of the Iowa Boys' Training School as well as all of the other teachers and students whose co-operation made this work possible.

Both authors were supported in this work by Ralph Drake Perry Fellowships.

work into sharper focus and permit a more integrated and sophisticated research program in the future. Two of these 10 reports have been published (**4**, **5**). The present report concerns nine studies which were carried out in 1957-58, none of which reached the level of significance. In view of the experimental variations introduced from our original design, there seems to be justification in presenting the results of these individual studies, in abbreviated form, for the inherent interest the reader may find in them. Also, we are, in this one report, acquainting the reader with a complete summary of our work to date on the relationship of ESP scoring and teacher-pupil attitudes. Although the studies appear collectively in the tables, we are not presenting our data for the purpose of giving a statistical whole but rather for what they may independently contribute to ESP research.

Furthermore, we have been accumulating data on the relation of ESP scoring and student academic achievement as expressed by the class grade the student is receiving in the class in which the ESP test is conducted. We wish, therefore, to present this full account of all our work preparatory to a later publication of grades based in part upon it.

Procedure

The method of testing used in all of the experiments summarized here is essentially the same as that used in our original work. This procedure has been written out in detail in previously published reports (**1**, **2**), and therefore only the main points will be mentioned here.

All materials used for testing were prepared at the Parapsychology Laboratory of Duke University and mailed to the teachers by the authors. Each teacher received a standard sheet of instructions with which she was asked to become familiar before informing the students of the nature of the experiment (see A-W 1).

The teacher was asked to distribute to each of the students a manila envelope which contained a random order of 125 ESP symbols (star, square, cross, wavy lines, circle) written on a standard ESP record sheet which in turn was enclosed between two opaque sheets of paper. The manila envelope was stapled shut. Stapled

on the outside of it were two blank ESP record sheets with a piece of carbon paper between them.

Each student was instructed to write on the outer record sheet his guess as to the order of the ESP symbols which appeared on the record sheet enclosed.

Attached to the back of each manila envelope was a smaller white one, inside of which was a questionnaire designed to assess the student's attitude toward the teacher giving the test. Each student was requested to answer the questions on this sheet, then put it back in the envelope and seal it, thereby assuring himself that the teacher would not be able to read the contents.

An important part of the procedure was that the teacher, when giving the directions, asked the students to do their best not only for themselves, but also for the teacher. It was hoped that this would serve to point up the interpersonal relationship between teacher and pupil which all of these experiments were designed to investigate.

The students were acquainted with the five ESP symbols and were requested to make their guesses at their own speed. When everyone was finished, the teacher collected the envelopes.

The teacher was also asked to prepare a master list of the names of the students in the class as well as the academic grade each was receiving in the class at the time of the ESP test. On this same sheet the teacher answered "yes" or "no" to the question: "Would you include this student in an ideal class if you had the opportunity to form one?" All of the material was then mailed to M.A. at the Parapsychology Laboratory. It was thought that a question stated in this manner, although it did not directly ask whether or not the teacher liked the student, would nevertheless get at this fact indirectly. The teacher, it was assumed, would have to consider all of the various factors making up this attitude toward the student in order to answer the question.

When the test materials were received, each experimenter drew up her own master sheet working from the list the teacher had sent. When this was done, each had in her possession a sheet of paper on which were written the name of each student, his academic grade, and the rating given him by the teacher.

Next the experimenters opened the white envelopes which con-

tained the student's answers to the questions about their teacher. Each experimenter, independently of the other, scored these questionnaires according to a system of rating which had been decided upon prior to the experiment. The student ratings were entered on each experimenter's master sheet. When this task had been completed, the experimenters compared the entries on their master sheets. Where discrepancies were noted, each experimenter referred to the original records until they were in agreement.

The manila envelopes were opened next. Since the authors had retained a carbon copy of the target sheet at the Laboratory, they now possessed two copies of both the student's guesses and the target symbols. Each experimenter checked a different set of calls and targets independently of the other. These scores were entered on the master sheets and the two orders of results were compared, run by run. If a disagreement in score was found, each went back to the score sheets and checked again until the mistake was found, agreed upon, and corrected. Each experimenter then totalled the results on the basis of attitude classification and compared her results with the other until the correct figures were agreed upon.

The ESP data were later rechecked, call by call, by Mrs. William Stober or Mrs. James Moser of the Parapsychology Laboratory. When additional errors were found, the authors once more independently returned to the data and, if necessary, rechecked their analyses accordingly.

A word might be said here with regard to the type of questionnaire that was used to rate the student's attitude toward his teacher. With but two exceptions, the questionnaires used were identical with those which have already been published in connection with our original work. The high-school questionnaire can be found in reference 2, page 156. The questionnaire used in all tests at the fourth-, fifth-, and sixth-grade level is printed in full in reference 1, page 12. When the A-W test was administered to the seventh and eighth grades, the high-school questionnaire was used. The two exceptions were the Boys' Training School experiment and the two college classes of Miss Sally Springer. Further details about these

questionnaires are given in the reports dealing specifically with these experiments.

No matter what questionnaires were used in any of these experiments, they were broken down into three attitude categories corresponding to the 10-9 range or highly positive group, the 8-7 range or moderately positive group, and the 6-0 range or negative group. When "total positive" is referred to in the following reports, it means the combination of the highly positive and moderately positive attitudes, or 10-7.

The only other variation in procedure concerns the two precognition series which were carried out in a school in San Diego with the tenth grade and in Wheaton with the fifth and sixth. These experiments were done exactly as all the others except for the following differences. Instead of manila envelopes, the students were given two plain ESP record sheets which were stapled together with a piece of carbon paper inserted between. In the San Diego experiment, a white envelope containing the student questionnaire was stapled to the back of the record sheets. In the case of the Wheaton fifth and sixth grades, the principal administered the questionnaire to the children before the test was given. Then she mailed them to M.A. at the Parapsychology Laboratory.

The students in both experiments were instructed to write in the Call column of the record sheet the ESP symbol that they thought would appear in the corresponding position in the Card column in the future when the runs were checked. Before the tests had been mailed to the teachers from the Parapsychology Laboratory, the record sheets which were to be used for the guesses had been numbered. When the test materials were returned, a target entry in the random number table was determined by the standard method currently in use at the Parapsychology Laboratory. For further details concerning this method, see reference 6, page 38. Beginning with the first target sheet which had been filled in, each sheet was numbered numerically to correspond with the numbers on the sheets with the subjects' guesses. The ESP target symbols were written in the Card columns on the target sheets, and carbons were made. M.A. and R.W. then independently matched the calls with the targets. The data were analyzed and rechecked

from this point on in the manner which has already been explained in detail.

The figures which bear on the hypothesis that there is a relationship between level of ESP scoring and teacher-pupil attitudes consist of the number of students in each attitude classification who scored above or below mean chance expectancy, (M.C.E.). It was the rule in the first three published articles (3) to include the scores that fell exactly at M.C.E. with those scores that fell below M.C.E. Since there is neither a logical nor a statistical necessity for this, we have in this survey deleted all scores at M.C.E. Where the results are given from the previously published work in the A-W series, the figures are presented in this new manner so that there will be a common basis for comparison.

Boys' Training School Test

This study was carried out in a training school (reform school) in Iowa for boys between the ages of ten and eighteen. The school program is organized to follow the standards set by other accredited public schools. Four academic teachers, two men and two women, conducted the A-W clairvoyance test with 69 boys. The only alteration from the original test was in the student questionnaire, the wording of which was adjusted to meet the environmental situation. However, the point value of the questionnaire ranged from 0-10 to coincide with the original questionnaire.

The 69 subjects did a total of 345 runs with a score of 1733, which is 8 above M.C.E. The teachers were positive in attitude to 43 subjects, whose average run score (A.R.S.) was 5.11. The 26 subjects for whom the teachers registered a negative attitude gave an A.R.S. of 4.89 (see Table 1).

Thirty-four subjects, in the highest attitudinal category (10-9), had an A.R.S. of 5.10, whereas the 16 subjects who were moderately positive in attitude (8-7), gave an A.R.S. of 4.85. The 19 subjects negative in their feeling for the teacher had an A.R.S. of 5.04 (see Table 2).

The rate of ESP scoring with the interacting attitudes of students and teachers is shown in Table 3. Twenty-four subjects, who were very positive in their feeling for the teacher and who were

liked by the teacher, gave an A.R.S. of 5.18. The teachers liked nine subjects who were moderately positive in their feeling for the teacher and the A.R.S. was 5.07. The 33 subjects with mutually positive attitudes gave an A.R.S. of 5.15. Mutually negative attitudes were expressed in nine cases and the A.R.S. was 5.13.

In those parts of the three tables which present the number of subjects scoring above and below M.C.E. it will be noted that the direction of scoring is in the expected manner in all three cases. Although subjects with negative attitudes and the negative interacting attitudes of teacher and student show a positive A.R.S., there are more scores below M.C.E. than above in both of these instances. Eleven of the 19 students who expressed negative attitudes scored below M.C.E. while only seven scored above.

One might suppose that the environmental situation of a "training school," even in academic classes, would be somewhat restrictive to an expression of attitudes, as the interpersonal relation between teacher and student would probably not be comparable to that found in a public high school. While this is undoubtedly at least a partial truth, it is all the more interesting that the scoring trends should be similar to our original findings.

Hillsboro Test

Four teachers from a high school in Hillsboro, North Carolina, each gave the A-W test to one of her classes. The teachers had previously administered ESP tests, and most of the students were also familiar with envelope clairvoyance experiments. The procedure in this experiment was identical to that followed in the original A-W work.

A total of 105 students obtained a score of 2631 on 525 runs, with a deviation of +6 and an A.R.S. of 5.01.

Seventy students liked by the teacher had an A.R.S. of 5.01. Thirty-four students disliked by the teacher obtained an A.R.S. of 5.02. There is no teacher rating for one of the students (see Table 1).

From the students' point of view, 48 placed the teacher in the highly positive category and the A.R.S. here was 4.99. Thirty-six students rated the teacher 8-7 and obtained an A.R.S. of 5.10. The

A.R.S. for the 84 students who rated the teacher positively (10-7) is therefore 5.04. Twenty students indicated that they did not like the teacher and their A.R.S. was 4.93. Again, one student did not complete a questionnaire about the teacher (see Table 2).

From the point of view of mutual attitudes, 36 students who were liked by the teacher placed her in the highly positive 10-9 category where the A.R.S. was 4.89. Twenty-five students liked by the teacher placed her in the moderately positive 8-7 category and here the A.R.S. was 5.09. The combined A.R.S. for the 61 cases with mutually positive attitudes was therefore 4.97. Eleven students disliked by the teacher also did not like her and the A.R.S. was 4.62 (see Table 3).

When the students' scores in all attitude categories are divided on the basis of above and below M.C.E., 32 who were liked by the teacher scored above, while 34 scored below. Of the 34 students that the teacher disliked, 15 scored above M.C.E. and 16 below.

Nineteen students who rated the teacher 10-9 scored positively, while 25 obtained negative scores. In the 8-7 category, 22 students scored above M.C.E. and 15 below. Of the 20 students who disliked the teacher, six scored positively and 10 scored negatively.

Where the teacher-student feeling was mutual and the students rated the teacher 10-9, 13 scored above while 20 were below M.C.E. Where the teacher liked the student and they rated her 8-7, 15 scores were above M.C.E. while 11 fell below. Where neither student nor teacher liked each other, there were two scores above M.C.E. and 11 below.

None of these breakdowns is significant when evaluated by chi-square, but it may be of interest to note that in all three categories (teacher, student, and the interaction between the two) the positive attitudes generally are in the reverse direction from that expected, whereas in the case of the negative attitudes, in all three cases there are more scores below M.C.E. than above.

It may be that previous experience with the test situation cut the edge on the separation of ESP scores on the basis of attitudes in this experiment. There is a need for more work in this area to determine the effect of the novelty of the test and whether this has any bearing on the effectiveness of attitude breakdowns.

Although the over-all deviation for this experiment was only +6, there was considerable variation among the different classes. Three teachers had positive class deviations of 15, 25, and 30, respectively, with a total of +70 on 380 runs. Cutting across these results was a negative deviation of −64 on 145 runs, which yields a significant CR of 2.66 in the fourth class. The combination of the chi-squares for the four classes is 11.92 (4 d.f.) with a probability of P=.02. The first three deviations are rather in line with what the authors have found in classroom tests, but the deviation in the fourth class once more raises the point that in the design and evaluation of future experiments, more attention should be given to the individual teacher and class in an attempt to determine the factors involved which lead to psi-missing.

San Diego Precognition Test

An English teacher at San Diego in the California school system wrote A-W of his interest in the school work. He expressed a desire to conduct some tests with his tenth-grade classes. Following an exchange of letters, the teacher and the authors agreed that here was an opportunity to carry out the A-W experiment, using a precognition rather than a clairvoyance test. Details of the procedure have already been discussed.

The teacher administered the test to three of his classes on one day and to the other two on the following day. There were 124 subjects, who did 1240 runs, or 31,000 trials. The over-all score of 6165 gives a deviation of −35 from M.C.E. The combination of the chi-squares of the five classes is 10.63 (5 d.f.) which gives a probability of P=.05.

Two students did not complete their questionnaires, and the teacher did not feel well enough acquainted with three students to comment on his feeling toward them. This accounts for what might seem to be discrepancies in the following presentation. There were 86 subjects whom the teacher liked and who gave an A.R.S. of 5.03 as compared to an A.R.S. of 4.81 for the 35 subjects who were not liked by the teacher (see Table 1). With the liked group 41 scores were positive whereas 36 were negative. Of the 35

subjects not liked by the teacher, 10 scored positively as compared to 22 who scored negatively.

Thirty subjects, highly positive in their feeling for the teacher, gave an A.R.S. of 5.01, while 33 subjects, moderately positive in feeling, had an A.R.S. of 4.95 (see Table 2). The 63 positive subjects had an A.R.S. of 4.98. There were 27 positive scores and 29 negative scores. The 59 subjects who did not like the teacher gave an A.R.S. of 4.98. Of these scores 26 were positive and 27 negative.

Mutually positive attitudes were expressed in 48 cases, and the A.R.S. was 5.02 for these subjects (see Table 3). There were 22 positive scores while 21 were negative. In the group who did not like the teacher and were not liked by him, there were 22 who had an A.R.S. of 4.86. Seven scores were positive and 13 negative.

We had observed that most of the teachers in their first two high-school experiments chose, naturally one would suppose, one of their better classes to test. In view of this, M.A. wrote to the teacher, before the results were returned to him, inquiring as to which of the five classes he considered his best. He replied that his first class had this distinction. The A.R.S. was 5.28 and the over-all deviation in this class was +72 on 260 runs, with a CR of 2.23 and a suggestive probability of P=.02. The 12 subjects in the mutually positive classification gave an A.R.S. of 5.47. There were no mutually negative cases. The A.R.S. in class two was 4.73; in class three, 4.98; in class four, 4.93; and in class five, 4.95.

This study suggests that a further investigation might be in order. Using the original A-W design, that is, having one teacher test only one of his classes, significant results might be anticipated from a precognition test as well as a clairvoyance test.

Missouri Experiment

A teacher of mathematics, Miss Sally Springer, at a junior college in Missouri visited the Parapsychology Laboratory in the summer of 1957. She met with M.A. and together they reviewed the ESP work in the classroom. This teacher was eager to try the A-W test on some of her own classes.

Accordingly, in November of the fall semester, she gave the test to three of her classes. The first test was given to a senior class in high-school mathematics. The results are therefore broken down into two categories: high school and college.

The only distinction between the test procedure used in this experiment and that followed in the usual A-W work is that a different questionnaire, geared to the college level, was used by the students to rate their teacher. This scale consisted of points between 0 and 10 and was broken down into the three standard attitudinal categories.

In the high-school class, 24 subjects did five runs each, giving a total score of 597. This gave a negative deviation of 3 and an A.R.S. of 4.98.

In the attitude breakdown, one student did not receive a rating by the teacher. The teacher was positive toward 16 subjects who had an A.R.S. of 5.08. The teacher was negative toward seven students who gave an A.R.S. of 4.66 (see Table 1).

From the student's point of view, 16 placed the teacher in the highly positive, 10-9 category and they gave an A.R.S. of 4.85. Five students put the teacher in the 8-7 or moderately positive category and their A.R.S. was 5.40. The A.R.S. for the total positive student attitude group is 4.98. There were three students who disliked the teacher and their A.R.S. was 4.95 (see Table 2).

When the attitudes of teacher and students are combined, where the students rated the teacher 10-9 and she liked them, the A.R.S. was 5.02. Two students liked by the teacher placed her in the moderately positive category and the A.R.S. was 5.60. The total positive student attitude average was therefore 5.09 where the teacher liked the student. When both teacher and students disliked each other, the two students involved had an A.R.S. of 4.90 (see Table 3).

In order to evaluate the relationship between attitude and level of ESP scoring, the number of students in each attitude classification who scored above or below M.C.E. were counted. Although none of these figures reaches significance by the chi-square evaluation, in all three categories (teacher, student, and mutual attitudes) there are more students with positive attitudes scoring above M.C.E.

than there are students who score below. Where negative attitudes are involved, more students score below chance than above.

* * * *

In the college classes, there were 63 subjects doing five runs each who scored 1595 on 315 runs with an average run score of 5.06. The teacher liked 50 students whose A.R.S. is 5.14. Thirteen students were disliked by the teacher and the A.R.S. is 4.75 (see Table 1).

With the student attitudes, 38 rated the teacher 10-9 and the A.R.S. was 5.02. The teacher was given a moderately positive rating by 21 students, and the A.R.S. was 5.07. The total positive student attitude group contains 59 subjects whose A.R.S. was 5.04. Four students were negative in attitude toward the teacher and the A.R.S. was 5.45 (see Table 2).

The A.R.S. where the teacher liked the 31 students who placed her in the highly positive category was 5.11. She was positive toward 16 students who were moderately positive toward her and who averaged 5.20. At the other end of the scale, only one student the teacher disliked also rated her in the negative category. The A.R.S. was 6.20 (see Table 3).

When the number of students scoring above change in each attitude category are counted, in all three tables it can be seen that on the basis of the attitudes of the teacher, the student, and the interaction between them, there were more scores above M.C.E. than below. Where the attitudes are negative, only in the case of the teacher's attitude are the results in the expected direction. In both the student and the interaction categories, there are more students scoring above M.C.E. than below.

One of the most interesting points that can be made about this study has come out of the comments that the teacher made about the classes before any of the scoring had yet been done. She had this to say about the first class she tested: "This class is noisy and hard to control . . . they were *not* quiet as they should have been while taking the test. Frankly, I think I did a lousy job of presenting it to them." The 24 students in this class scored 596 on 120 runs, which gives a negative deviation of 4.

About the second class she said: "Class very cooperative. I did a better job of presenting the material." In this case the 25

students participating did 125 runs and scored 667, which gives a positive deviation of 42.

In testing the third class, she says: "This class was cooperative, although a little noisy. *They had just received grade cards before the test.*" Here the 38 students scored 928 on 190 runs, which yields a negative deviation of 22. There have been some indications in this and other work that it may not be the optimum time for giving an ESP test within the class situation when there are important disruptive events in the ordinary school schedule, such as vacations, or the distribution of grade cards. The results of this third class may be a further indication of this. However, these *ex post facto* speculations should be verified by further research.

On the whole, the results of this experiment indicate that it is worth while and possible to extend the A-W clairvoyance test through the junior college level. It should be pointed out that the only other difference between this experiment and the original A-W work is that in the former the same teacher tested three classes.

The St. Charles Test

Mrs. Merta Reed of St. Charles, Michigan, presented an opportunity for the authors to extend the original A-W tests to include the seventh and eighth grades. Mrs. Reed, an experienced teacher of many years in the St. Charles school system, met with the two seventh- and two eighth-grade teachers prior to the day of testing. The four classes were then tested on the same day and at the same hour. There were 112 subjects, who did a total of 560 runs, or 14,000 calls. The total number of hits was 2887, which is 87 above M.C.E.

Sixty-five subjects were given positive ratings by the teacher, and the A.R.S. for this group was 5.16. The 47 subjects to whom the teachers gave a negative rating also gave an A.R.S. of 5.16 (see Table 1).

The relationship of ESP scoring and student attitude toward the teacher is shown in Table 2. Fifty-four subjects rating the teacher highly positive (10-9) gave an A.R.S. of 5.14, while 40 subjects rating the teacher moderately positive (8-7) had an A.R.S. of 5.29. The A.R.S. for the 94 subjects who were positive in feeling was

5.21. The 18 subjects who did not like the teacher (6-0) gave an A.R.S. of 4.89.

The interaction of attitudes shows that 35 subjects who were very positive and were liked by the teacher gave an A.R.S. of 5.21, and the A.R.S. for the 19 subjects moderately positive in feeling and liked by the teacher was 5.20. The seven subjects with whom mutually negative attitudes prevailed had an A.R.S. of 4.89 (see Table 3).

The above- and below-M.C.E. scoring is in the expected direction in all of the tables except with regard to the students to whom the teacher had negative attitudes, where more positive than negative scores are recorded (see Table 1).

While this experiment did not quite reach the level of significance, it is clearly in line with the original findings with the exception of the positive scoring where the teacher is negative in attitude. It may be worth while to point out that the results in the eighth grade were in accord with the earlier finding in every respect. The CR for the two eighth-grade classes was suggestive at 2.11 with P=.02. It was the seventh-grade classes that gave the atypical scores.

Experiment with Non-academic Classes

This experiment involved six teachers in the non-academic areas of physical education, home economics, agriculture, and art. Two of the teachers were in the same Arizona school system; the other four were in schools in Arizona, Florida, Indiana, and Illinois. Exactly the same procedure as that used in the A-W 1 series was followed. There were a total of 161 subjects tested in these six classes and they did a total of 805 runs. The total number of hits was 4005, which is 22 below M.C.E.

The teachers indicated a positive attitude toward 93 subjects, who scored an A.R.S. of 4.98, whereas the 68 for whom the teachers registered a negative attitude gave an A.R.S. of 4.97 (see Table 1).

Sixty-six subjects rating the teacher highly positive (10-9) gave an A.R.S. of 5.01, while the 57 subjects moderately positive in feeling (8-7) scored right at M.C.E., 5.00. Expressing a nega-

tive attitude for the teacher, 38 subjects gave an A.R.S. of 4.87 (see Table 2).

Forty-two subjects who rated the teacher highly positive (10-9) and who were liked by the teacher gave an A.R.S. of 5.05. The 34 subjects liked by the teacher who were moderately positive (8-7) in feeling for the teacher gave an A.R.S. of 4.98. Thus in the 76 mutually positive relationships the A.R.S. is 5.00. Twenty-one subjects who were regarded negatively by the teacher and who rated the teacher negatively, had an A.R.S. of 4.89 (see Table 3).

As may be seen from the tables, there are more scores below M.C.E. than above M.C.E., regardless of positive or negative attitude.

There are many obvious differences in the structure and organization of academic and non-academic classes. One would suppose that a student might have a different attitude toward his, say, mathematics class than his physical education class. The former is more formal and intellectually challenging than the latter, which tends to be informal and playful. It is quite plausible that both teacher and student view each other differently in the two situations. However, the student questionnaire which was used in this study was the same as for the academic classes, and the results pose the question as to whether it is suited to the psychological environment of non-academic classes.

Wheaton Fifth- and Sixth-Grade Test

Mrs. Mary Davies, Principal of Hawthorne Elementary School in Wheaton, Illinois, came to the Parapsychology Laboratory of Duke University in the summer of 1956. She met M.A. and discussed with her the A-W experiment in the classroom, which had been recently completed. She was very enthusiastic both about parapsychology in general and, in particular, about trying to repeat the A-W experiment at the fifth- and sixth-grade levels. That fall she carried out an experiment in her own school. It yielded significant results and has already been published in this *Journal* (1). This experiment is part of what is now referred to as "the original A-W work."

In the summer of 1957, Mrs. Davies attended a parapsychologi-

eal workshop which was held at the Laboratory. At this time, as well as during a visit to Wheaton in September, 1957, by M.A. and R.W., plans were made for continued experimentation at Hawthorne Elementary during the subsequent school year.

In the first experiment at Wheaton, two fifth- and two sixth-grade teachers had participated. In this present experiment, the same sixth-grade teachers took part and one of the former fifth-grade teachers. Only these three teachers were used. It should be pointed out that the students in the sixth grade in this experiment had also participated in the ESP test at the fifth-grade level during the preceding year.

It has often been said that it is virtually impossible to "repeat" in all details an ESP experiment even with the same subjects and experimenters. If a person has already taken or given a test, he is, psychologically speaking, not quite the same person when he attempts to do it again by virtue of his prior experience. It has therefore been suggested that in order to approach the psychological conditions of the first experiment, one must introduce novelty into the second situation. In the experiment reported here, therefore, it was decided to try something these children had not done before, namely, precognition. Some suggestive evidence that precognition may occur in the classroom had previously been provided by the San Diego experiment, although this work had not been extended to the fifth- and sixth-grade levels. An attempt was made to do so in the present experiment.

The details of the procedure used in this experiment have already been given (see page 7). We need only point out that each subject did two sessions of two runs each for a total of four runs per student. Where there is a discrepancy in the tables between the number of subjects and the number of runs, it is due to the fact that not all students were able to be present for both sessions.

Ninety students did 358 runs with a total score of 1767 hits. This yields a negative deviation of −23, and the A.R.S. is 4.94.

Seventy-five students were liked by the teacher, and their A.R.S. was 4.90. The teacher disliked 14 students, whose A.R.S. was 5.24 (see Table 1).

From the student's point of view, 36 of them rated the teacher from 10-9 and achieved an A.R.S. of 4.70. Twenty-five students placed the teacher in the moderately positive 8-7 category and their A.R.S. was 5.06. The A.R.S. for the 61 total positive student attitudes is therefore 4.85. Twenty-nine students did not like the teacher and their A.R.S. was 5.14 (see Table 2).

In the interaction of the teacher and student attitudes, the teacher liked 30 students who placed her in the highly positive category. The A.R.S. was 4.60. Twenty-one students liked by the teacher were moderately positive in their attitude toward her. The A.R.S. is 5.01. Only four students disliked by the teacher also did not like her and their A.R.S. was 5.13 (see Table 3).

As to the number of students in each attitude classification who scored above or below chance, in all three categories where the attitudes are positive, approximately one-third more of the students go below M.C.E. than score above. In regard to the negative attitudes, although only a few subjects are involved, there are an equal number of scores above and below M.C.E. in the student attitude breakdown. In both the teacher and the teacher-pupil interaction, the number of scores above or below chance are slightly in the predicted direction. That is, more students in positive attitude classifications scored above M.C.E. than below, and more students in negative attitude classifications scored below M.C.E. than above.

Although the relationship between attitudes and level of ESP scoring in this experiment is not significant, there does seem to be a consistent trend. The results seem to go, in almost all cases, in the direction opposite from the evidence in the work done the previous year. It is hard to say what might be the cause of this, and at best one may only surmise. It is even impossible to consider the results merely from the point of view of alterations in psychological motivation, since the new factor of precognition is also involved.

Wheaton Fourth-Grade Test

It was the idea of Mrs. Mary Davies, the principal at Hawthorne Elementary School, to try another innovation in her school tests during the fall of 1957. The precognition experiment at the fifth-

and sixth-grade levels has already been described. The second innovation was an attempt to extend the A-W clairvoyance work to the fourth-grade level.

With two exceptions, the procedure followed was as nearly as possible the same as that used in the first experiment done at the fifth- and sixth-grade level the year before, which had yielded significant results. One exception was that Mrs. Davies introduced a change in the way the children were to think of and draw some of the ESP symbols so that this more closely approximated the usage familiar to children of elementary school age. It is the common practice to designate the square symbol by an L shape. To make this less confusing, the students in this experiment drew a square box. They were told to think of the "wave" symbol, which is written like an equal (=) sign, as a pair of bars such as they were used to playing on in the playground. The ESP star symbol is usually drawn as an inverted V. The children were therefore told to think of this symbol as a "Teepee" which in fact does more closely approximate the target symbol which was to be written on the record sheet inside the test envelopes. No changes were made in the use of the cross and circle symbols.

The only other variation from the standard procedure was that each child did one run a day for five days, which made a total of five runs per student. Any discrepancies noted between the number of runs and the number of students is due to the fact that not all students were present at all five sessions.

Eighty-four students did five runs apiece with a total score of 2018 on 409 runs, a deviation of −29, and an A.R.S. of 4.95. Sixty-seven students were liked by the teacher and had an A.R.S. of 4.92. The teacher disliked 13 students whose A.R.S. was 5.14 (see Table 1).

With regard to the student's attitude toward the teacher, 29 placed her in the highly positive 10-9 category and the A.R.S. was 4.71. Twenty-four students gave the teacher a moderately positive rating of 8-7 and their A.R.S. was 4.87. This means that the A.R.S. for the total positive student attitude was 4.78. Thirty-one students, on the other hand, gave the teacher a negative rating and their A.R.S. was 5.19 (see Table 2).

Where the teacher liked the student and the students placed the teacher in the 10-9 classification, the A.R.S. was 4.69 for 27 students. Eighteen students liked by the teacher gave her a rating from 8-7 and the A.R.S. was 5.00. Nine students, disliked by the teacher, also did not like her and the A.R.S. was 5.42. The CR of the difference between the total of this group with the highly positive attitudes toward the teacher who also liked them and the scores when the attitudes were mutually negative is 2.41, with P=.006. Here, as in the experiment with the fifth-and sixth-grades, there is a reversal in the trend from that which one might expect on the basis of the A-W work and also from the tests carried out at the same school during the previous year.

When we look more closely at the relationship between ESP scoring level and teacher-pupil attitudes by counting the number of students in each attitude classification who score above or below M.C.E., the same reversal of trend persists. As can be seen from the three tables, in all of the attitude categories there are more students on the positive attitude side who score below M.C.E. than above. With regard to the negative attitudes the reversal of this is true, and where one would expect more students to be scoring below M.C.E. they are instead scoring above. There is at present no explanation for this rather consistent trend of scoring negatively when a positive score was expected, but if this trend can be caught in future experiments it raises a question of great importance . . . that of a group effect apparently causing psi-missing. We shall return to this problem in a later paper.

Discussion

Many of the questions which have been prompted by these studies have already been discussed in their respective sections of the text. Of the nine studies reviewed, the first six were conducted in academic classrooms. In five of these six studies, the consistency figures are directionally similar to what we had found in our earlier work. The Hillsboro and the Wheaton precognition investigations, as has been pointed out, involved teachers and students who had previously taken part in a psi test. We do not say that the fact that the second test could not have the same novelty for

Table 1

ESP Results as Related to Teacher's Attitude Toward Student

Experiment	Teacher Positive					Teacher Negative				
	No. Subjects	No. Runs	Av. Run Score	No. ESP Scores		No. Subjects	No. Runs	Av. Run Score	No. ESP Scores	
				Positive	Negative				Positive	Negative
Boys' Training School	43	215	5.11	22	16	26	130	4.89	7	14
Hillsboro	70	350	5.01	32	34	34	170	5.20	15	16
San Diego	86	860	5.03	41	36	35	350	4.81	10	22
Missouri High School	16	80	5.08	9	5	7	35	4.66	1	4
Missouri College	50	250	5.14	24	19	13	65	4.75	4	5
St. Charles	65	325	5.16	34	27	47	235	5.16	27	18
Nonacademic	93	465	4.98	40	47	68	340	4.97	25	34
Wheaton Precog.	75	300	4.90	28	37	14	54	5.24	4	5
Wheaton Fourth	67	324	4.92	30	31	13	65	5.14	8	5
Total	565	3169	5.02	260	252	257	1444	4.96	101	123
A-W 1	132	660	5.22	69	49	96	480	4.86	40	49
A-W 2	121	605	5.26	67	46	84	420	4.85	32	43
Wheaton	80	316	5.09	38	35	33	130	4.72	9	22
Total	333	1581	5.21	174	130	213	1030	4.84	81	118
Over-all Total	898	4750	5.08	434	382	470	2474	4.91	182	241

the participants as the first was the reason for the changes in scoring by students in the different attitude classifications but only that this possibility should be investigated.

If we combine the chi-square results of the individual two-by-two contingency tables on scoring above and below mean chance expectancy, including those from the previously reported tests, we obtain a probability of .004 on the teacher attitude, .009 on the student attitude, and .00004 on the attitudinal interaction of student and teacher.

Our main interest in the tests reported in this paper, however, is due to the new experimental designs we may draw from them and to the basis which they provide for our studies of the relationship between ESP scoring and student class grade shortly to be published.

Table 2

ESP Results for the Three Levels of Student's Attitude Toward Teacher

Experiment	10-9 (Positive)		8-7 (Positive)		10-7 (Positive)		No. of ESP Scores		6-0 (Negative)		No. of ESP Scores	
	No. of Runs	Av. Run Score	No. of Runs	Av. Run Score	No. of Runs	Av. Run Score	Positive	Negative	No. of Runs	Av. Run Score	Positive	Negative
Boys' Training School	170	5.10	80	4.85	250	5.02	22	19	95	5.04	7	11
Hillsboro	240	4.99	180	5.10	420	5.04	41	40	100	4.93	6	10
San Diego	300	5.01	330	4.95	630	4.98	27	29	590	4.98	26	27
Missouri High School	80	4.85	25	5.40	105	4.98	11	8	15	4.95	0	1
Missouri College	190	5.02	105	5.07	295	5.04	25	23	20	5.45	3	1
St. Charles	270	5.14	200	5.29	470	5.21	53	37	90	4.89	8	8
Nonacademic	330	5.01	285	5.00	615	5.01	52	61	190	4.87	13	20
Wheaton Precog.	142	4.70	100	5.06	242	4.85	20	31	116	5.14	12	12
Wheaton Fourth	143	4.71	115	4.87	258	4.78	21	28	151	5.19	17	11
Total	1865	4.98	1420	5.04	3285	5.01	272	276	1367	5.00	92	101
A-W 1	515	5.35	330	5.09	845	5.25	92	61	295	4.56	17	37
A-W 2	420	5.16	330	5.33	750	5.23	80	60	275	4.73	19	33
Wheaton	168	5.25	140	4.99	308	5.12	36	36	138	4.66	11	21
Total	1103	5.26	800	5.17	1903	5.22	208	157	708	4.64	47	91
Over-all Total	2968	5.08	2220	5.09	5188	5.09	480	433	2075	4.88	139	192

While we could suggest many possible explanations for the results in each of these studies, we do not feel that a detailed interpretative discussion is justified at this time. We would like to consider the investigations in a broader context.

There have been a sufficient number of experimental studies to justify the conclusion that psi has been demonstrated in the classroom. There is reason to suppose that psi has functioned at all school levels, and that any test—GESP, clairvoyance, or precognition—could yield similar results. However, the tables in this report serve as reminders of the seeming necessity for structuring the experimental design in terms of the given school and class situation. We have now tested 49 classes. There has been enough

Table 3

ESP Results as Related to Attitude of Teacher and Student Toward Each Other

Experiment	Total Mutually Positive				Total Mutually Negative			
	No. of Runs	Av. Run Score	No. of ESP Scores		No. of Runs	Av. Run Score	No. of ESP Scores	
			Positive	Negative			Positive	Negative
Boys' Training School	165	5.15	18	11	45	5.13	3	6
Hillsboro	305	4.97	28	31	55	4.62	2	7
San Diego	480	5.02	22	21	220	4.86	7	13
Missouri High School	75	5.09	9	5	10	4.90	0	1
Missouri College	235	5.14	22	18	5	6.20	1	0
St. Charles	270	5.20	30	22	35	4.89	4	3
Nonacademic	380	5.00	34	39	105	4.89	7	12
Wheaton Precog.	204	4.78	17	27	16	5.13	1	2
Wheaton Fourth	218	4.83	19	23	45	5.42	6	3
Total	2332	5.02	199	197	536	4.93	31	47
A-W 1	550	5.37	64	34	185	4.63	12	22
A-W 2	505	5.30	57	38	175	4.53	9	25
Wheaton	230	5.15	28	25	52	4.27	1	11
Total	1285	5.31	149	97	412	4.54	22	58
Over-all Total	3617	5.11	348	294	948	4.76	53	105

variation among classes to indicate that, despite the over-all trend for the attitudinal breakdowns to be in the anticipated directions, the role of the teacher in the experiment is an important and complex one which must be studied further. Educational psychologists have long been interested in why a particular teacher with a particular class develops a congenial and productive atmosphere while the classes of others remain dull and colorless. There may be something for us to learn by considering this question.

It appears that the delicate nature of psi phenomena requires a psychological sensitivity from everyone concerned in the experiment, beginning with the experimenter. There are those who regard the experimenter as outside the experiment. These people have hastily accepted the procedures employed in the study of atoms or stars as models for studying man. If a competent experimenter repeats an experiment exactly, his results should be the same as those of any other investigator. Convenient as this would be for

workers in parapsychology, the strict repeatability of the findings of the exact sciences is not found in our area of investigation. This is not to say that it never will be, though we should remain open to the possibility that it may not. But one thing seems clear. In psychological experiments, you cannot regard the subjects as "other things out there." It has been stated many times in many ways that in psi research it is a psychological function that is being tested. The most obvious, and thus the least noticed, point is that a psi experiment is a social interplay between *all* of the personalities involved in the test. This should be emphasized to the point of repetitiveness.

None of these nine series was done as a logical extension of the A-W series. They were not, as A-W 2, projected as efforts to confirm the A-W 1. Rather, we were freely exploring in various directions for possible departures or variants of conceivable value. Hence the characterization of them as exploratory. Much of their value is in the fact that they did *not* yield better results or even as good as we had been getting.

References

1. Anderson, M. Clairvoyance and teacher-pupil attitudes in fifth and sixth grades. *J. Parapsychol.*, 1956, **21**, 1-11.
2. Anderson, M., and White, R. Teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1956, **20**, 141-57.
3. ———. A further investigation of teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1957, **21**, 81-97.
4. ———. ESP score level in relation to students' attitude toward teacher-agents acting simultaneously. *J. Parapsychol.*, 1958, **22**, 20-28.
5. ———. The relationship between changes in student attitudes and ESP scoring. *J. Parapsychol.*, 1958, **22**, 167-74.
6. Mangan, G. Evidence of displacement in a precognition test. *J. Parapsychol.*, 1955, **19**, 35-44.
7. van Busschbach, J. G. An investigation of extrasensory perception in school children. *J. Parapsychol.*, 1955, **17**, 210-14.
8. ———. A further report on an investigation of ESP in school children. *J. Parapsychol.*, 1955, **29**, 73-81.
9. ———. An investigation of ESP between teacher and pupils in American Schools. *J. Parapsychol.*, 1956, **20**, 71-80.

Parapsychology Laboratory
Duke University
Durham, North Carolina

PRELIMINARY EXPERIMENTS ON PSI PHENOMENA WITH RELATIVELY UNCONSCIOUS PARTICIPATION

By Ronald E. Shor, Laurence S. Youell, and Hershel Berkowitz

ABSTRACT: This paper contains an outline of the logic and experimental design of a series of projected experiments and the results obtained on the first, exploratory series. The general objective of the work is to test ESP in the laboratory under conditions which favor spontaneity by minimizing the operation of the higher mental processes and by increasing motivation to a point more nearly equal to that of real-life situations. Basically, the electrical apparatus for the tests requires the subject to respond, sometimes sensorially and sometimes extrasensorially, to the flashing of a light. In the preliminary series reported here, only stimuli calling for extrasensory responses were used. Although statistical significance was achieved, the authors point out that technical and methodological difficulties arose which throw some doubt on the findings.

Mr. Shor is a Public Health Service Research Fellow who is doing research at the Massachusetts Mental Health Center. He also teaches psychology at Brandeis University. Mr. Youell and Mr. Berkowitz were undergraduates at Brandeis when they worked on the experiment.—Ed.

Introduction

Although psi has been repeatedly demonstrated in the laboratory, the evidence is almost always fragile and trifling. Perhaps no one would pay any attention to it (or indeed have discovered it at all) if it had not been fathered by the more potent phenomenon of psi which occurs spontaneously in natural life circumstances. These instances of spontaneous psi are frequently claimed to occur with great detail, accuracy, and clarity. If we take as our working assumption that both meager laboratory psi and strong spontaneous psi can exist, then we are left with the paradox of why a phenomenon can exist in a sharp and clear form in natural life circumstances and be only dimly perceivable in the rigorous confines of the laboratory. The need to bridge the gulf between these two situations and the related issue of achieving repeatability are the great problems of parapsychology today. The dragon of unrespectability is mostly slain and there is little need

to be overly concerned with simple "proof" of psi, but the impasse with regard to strength and repeatability of the phenomenon remains almost as formidable as before.

In trying to face this problem, a question has repeatedly struck many investigators as crucial, to wit: in what ways does the total situation in which spontaneous psi occurs differ from the total situation in which psi is investigated in the laboratory? In reviewing the parapsychological literature, many such differences are observable. Four of the major differences are described below. They are discussed with an eye toward bringing some of the features of spontaneous psi into the laboratory.

I. In laboratory psi the motivation to succeed is either a simple desire to cooperate or it is due to task- or ego-involvement, competition, or token rewards—all of which are extraneous or peripheral to the vital needs of the subject. Spontaneous psi occurs more frequently when the person is vitally concerned with the event. While it is not possible to make success a "life and death" matter with humans in the laboratory, a strong motivation to avoid errors could be utilized, such as electric shock. Shock might tend to tap a more salient range of organismic purposes.[1]

II. The form of response in laboratory psi is typically in terms of higher mental processes such as: guessing cards, influencing dice, reproducing pictures, and so forth. In spontaneous psi the experience often begins as a "feeling" or sudden nonverbal experience of a total state of affairs. It is hard to implement such a thing in a laboratory but perhaps one subordinate feature of it can be achieved; that is, the use of simple nonverbal motor responses. Tyrrell investigated this supposition with encouraging results (**7**).

III. In laboratory psi, a modest time interval is usually allotted to make judgments. This time interval is given with the understanding that some kind of response is eventually required. Spontaneous psi, by definition, comes without warning. Perhaps decision time allows mobilization of other mental processes which becloud psi and lower its efficiency in the laboratory. By requiring an instantaneous

[1] See also in this regard the work of McElroy and Brown (**1**), who utilized a shock feedback of errors both to motivate subjects and to supply information of successes and errors that subjects might learn to discriminate extrasensory cues.

response (on the order of a reaction time) psi might occur in a less obscured fashion. Also, perhaps, if the subject's task could be made painfully simple, repetitious, monotonous, and boring it may be that fewer extraneous mental processes would interfere.[2]

IV. In laboratory psi, the subject may decide how he will respond on the next trial *before* the stimulus to judge is given. It is possible, then, that subjects tap predecisions rather than psi responses when required to respond. (Predecisions may themselves have their own psi processes but these may confuse the issue.) There is no way to get rid of this hypothetical problem entirely, but there may be a way to minimize it. Both boredom and monotony may help achieve this end. Also, it might help if the subject were forced to remain constantly "set" to respond to real sensory information. That is, trials requiring sensory discriminations can be randomly intermixed with trials requiring extrasensory or guesswork discriminations. If allotted decision time were then preset so that only enough reaction time were available to make a response *if* the subject were "set" on a sensory basis (but not enough time if he were "set" in terms of a predecision) perhaps pretrial decisions would be minimized on extrasensory trials also.

All four of these points resolve into three basic assumptions: that psi is an unconscious, primary process; that secondary conscious processes interfere with it in various ways; and that it is called forth in clearest form when vital motives are involved. Spontaneous psi differs from laboratory psi in that spontaneous psi is less consciously directed and refers to vital concerns of the person. Laboratory psi is more confounded by conscious processes and is less vitally motivated. These conclusions are not, of course, unique with the writers. The unconscious source of psi has been alluded to so widely that references are superfluous.

It had been our aim to design an experiment in which relatively unconscious responses could be elicited from subjects and motives somewhat more vital than in the usual experiment could be involved. To accomplish this end, the eight procedures outlined above have

[2] Experimental evidence exists to support the opposite supposition: that boredom and fatigue debilitate psi. (**2**; **3**; **5**, pp. 126-31; **6**). Too little is known, however, to take the evidence as conclusive.

been mobilized; to summarize: (*a*) feedback of information about errors via shock, (*b*) simple motor responses, (*c*) instantaneous responses, (*d*) boredom, (*e*) repetitiousness, (*f*) monotony, (*g*) utter simplicity, and (*h*) sensory "set." An experimental model has been conceived to investigate systematically all of these factors or varying combinations of them as desired. The complexity of our plans made it necessary to design an apparatus which included in its construction all of these varied features. The apparatus was constructed so as to be completely automatic, precisely timed, and yet highly flexible. Even recording of scores was made automatic with the hope of eliminating the dread possibility of errors due to experimenter's bias.[3]

Apparatus

The basic design of the experiment is to present three white lights to the subject's view. Those on the left and right call for sensory responses. That is, when the left-hand light comes on, the subject is to press a pushbutton underneath that light, and when the right-hand light comes on, he is to press a button on the right. The central light calls for an extrasensory response. The subject, in this case, must decide between the left- and right-hand buttons, the correct one having been determined electrically.

The equipment is in two parts, connected by a multi-conductor wire. Half of the equipment is located in an experimental room on the second floor of one of the psychological laboratories at Brandeis University. It is here that the subject performs his task. The remainder of the equipment is located in a separate room in the basement of the same building, and it is monitored by the experimenter.

The equipment in the basement consists of (*a*) a programmer which automatically advances a prepunched tape, (*b*) related equipment to translate the information on the tape into the appropriate experimental sequences, (*c*) a repeating dual-cycle timer to control signal intervals and rest intervals, and (*d*) a D.C. power supply unit to supply current for time delayed relays and non-arcing circuitry in the programmer.

[3] For other parapsychological investigations where the design is automatized see Webster (**8**), Tyrrell (**7**) and Seeley (**4**, p. 344).

On the second floor, there are three white lights mounted horizontally to the subject's chair. In front of the two outermost lights are two pushbuttons which the subject can selectively depress. There is also a small pilot light to report on errors, an inductorium and wrist strap to give shocks when errors are made, a regular pilot light, and a toggle switch which starts and stops the whole apparatus.[4]

Except for the D.C. power supply which has its own separate controls, a single toggle switch feeds voltage to all of the equipment. When this circuit is closed, a red pilot light turns on. When the prepunched tape is put into the automatic programmer, a microswitch mounted on the programmer is closed by the tape, and this extinguishes the red light, turns on a green light, and feeds voltage to the body of the equipment. After a few seconds, the repeating dual-cycle timers begin to function, and the subject can set the whole apparatus in motion by closing the toggle switch mounted on his part of the equipment. According to the prearranged pattern of holes on the tape, one of the three white lights will come on for the time interval preset on the first timer. One of four impulse counters will record this occurrence. An extra counter will also tally it.[5] At the end of the time interval, the white light will go off; and if the subject has not depressed the *one proper* pushbutton, an error is recorded on one of four other impulse counters. Coincidentally with the recording of an error, the subject may receive feedback information about the error in the form of the flashing of the error light or as an electric shock through an inductorium with electrodes strapped to his wrist. After the error is recorded, the solonoid attached to the programmer fires twice, advancing the tape to its next position. During each trial, the prepunched tape makes or breaks two independent circuits, thus yielding four possible alternatives per trial. These are: (*a*) right white light, right-hand response correct; (*b*) left white light, left-hand response correct; (*c*) center white light, right-hand response correct; and (*d*) center

[4] The authors will be glad to provide interested readers with a copy of the circuit diagram.

[5] There would be no difficulty in including a multichannel recorder for a continuous record of responses over time. This has not been done as yet, however.

white light, left-hand response correct. A complex of simple relays, time delayed relays, and interlocking relays translate these patterns on the tape into their appropriate electromechanical expression. After another time interval, a new cycle begins. At the end of the tape the microswitch again opens, stopping everything.

Procedure

The procedure discussed in this section is typical for the projected series of experiments as a whole. The procedure of the one experiment reported in this paper differs from the general procedure in a number of minor ways which will be noted.

Subjects are shown the part of the apparatus located in the basement room, and the experimental design is explained to them. Frequencies on the impulse counters are recorded in duplicate on specially prepared sheets with carbon paper between. They are carefully checked by the subject and experimenter. The subject keeps all of these papers in his possession throughout the experiment.

Subjects are then led to the part of the apparatus on the second floor, and their task is explained. They are told that each time a light comes on they must depress one of the two pushbuttons and keep it depressed until the light goes out. They are then to release the button and await a new trial. They are informed that failure to press either pushbutton, pressing both at once, pressing the wrong one, or releasing the pushbutton prematurely results in an automatic error. Only if they press the one correct pushbutton and keep it depressed until the white light goes out can they escape an error.[6] They are told that they themselves are to start the experiment by closing the toggle switch on the side of the machine. They are warned to depress one pushbutton before closing the toggle switch so as to have a choice for the first trial which will follow immediately thereafter.

Once the switch is closed the subject must press one of the pushbuttons each time one of the three white lights comes on. When

[6] These statements are all factually correct. Because there are so many ways to make an error, deviations in the negative direction are inapplicable to the psi hypothesis. Deliberately, then, an experimental situation has been constructed which is insensitive to the usual subtleties of psi phenomena, for the sake of pursuing a special kind of effect.

one of the outer white lights comes on, his selection can be made on a simple sensory basis; he must press the left pushbutton when the left white light comes on and the right pushbutton when the right white light comes on. When the center light comes on, however, there is no sensory basis for selection. Half of the time the left pushbutton is appropriate and half of the time the right. The subject must make his choice on the basis of guesswork and/or psi processes.

In experiments where the subject is to be given a limited amount of reaction time, a preliminary test is devoted to the determination of the appropriate interval for each person.

At the end of the experiment, when the apparatus automatically shuts itself off, the subject returns to the basement room and aids the experimenter in recording and checking the new totals on the impulse counters in duplicate on the specially prepared sheets. At the bottom of these sheets the subject is asked to sign the following statement. "I am completely satisfied that the frequencies reported above are without any errors in copying from the counters." The two copies are then separated and the subject is given the carbon copy to keep permanently.

Present Experiment

One experiment in our projected series has been completed so far. In many ways it is the simplest, most basic experiment in the series and does not include all of the special characteristics noted above. In this experiment only the center light was used so that the factor of sensory "set" was not involved. Modest time intervals were given both to react to the white signal lights (two seconds) and for the rest intervals between trials (two seconds). Thus the factor of instantaneous responses was not mobilized. Moreover, no feedback of information whatsoever (neither light nor shock) was given. The factors of boredom, repetitiousness, monotony, and utter simplicity remain, however. Informal reports from subjects bear this out. Many subjects report that they quickly tended to lose conscious interest in the task, and that there remained a kind of passive compliance, with little active choice. Because the flashing white light is bright enough to see by peripheral vision, subjects would sometimes stare out of the window or look

about the room. One subject reports that his mind wandered for a brief interlude (about ten minutes) to other thoughts and that during this interlude he completely forgot the task. He continued it automatically, however, and achieved a high score.

Subjects

Approximately 25 volunteer undergraduate and graduate students from Brandeis University (and one undergraduate from Harvard) were run. Of this number, the records of all but eight subjects were excluded from consideration. Data were accepted or rejected on the basis of mechanical excellence alone and *never* on the basis of the results themselves. A number of unexpected technical difficulties were encountered which made it inadvisable to use the other records. Most of the difficulties were a direct result of possessing equipment which was not always of heavy-duty construction. No problem of this kind was found, however, which cannot be surmounted by improved construction.

Results

The findings are summarized in the table below.

GENERAL RESULTS

Subjects	No. Hits	Percentage of Success
L.Y.	418	54
P.C.	405	52
E.C.	397	51
R.B.	430	55
S.C.	383	49
R.L.	415	54
C.K.	378	49
R.S.	404	52
Average	404	52

Total trials = 6224; total hits = 3230; Dev. = +118; CR = 3.00; P = .0013

Each subject made 778 trials during a running time of approximately 52 minutes. It is to be expected that a score of 50% correct (or 389) would be achieved by chance. The CR of 3.00 on all scores combined yields a highly significant difference (P=.0013).

Before the experiment, subjects were asked a series of questions to evoke their attitudes and beliefs in regard to telepathy, psychoki-

nesis, clairvoyance, and precognition. They were also asked if they liked mechanical gadgets and to what extent they were interested in the experiment. Because of the small number of subjects who participated, it is not possible to establish a relation between their scores and their attitudes on the basis of this first experiment alone.

Because our impulse counters were not of the best mechanical quality, a chattering of the counters occurred occasionally, so that a few extra digits were recorded on those counters that tallied the targets selected by the machine. Since the actual subtotals for the four possible choices are known, it was these that were used to derive the number of successes, by subtracting the errors from the number of times the target came up. Tallies on the error counters were not changed, however, even where it was known that they were a few digits too high, in order to maintain full conservatism. The inaccuracy of the counters was not of a type to create the suspicion that the opposite might have occurred; i.e., that the counters might have failed to record a digit. The very remote but nevertheless distressing methodological possibility exists, however, that this theoretically could have occurred, weakening considerably the force of our results.

Discussion

Along with rebuilding the apparatus with all heavy-duty equipment, the methodological inadequacy in the apparatus which is noted above must be corrected before further work is undertaken. It is essential that an *independent* tally of successes be arrived at. It is poor design to derive successes by subtracting errors from totals, as is now the case. Successes must be tallied on another set of four counters, engineered so that safeguards offset even a remote possibility of spuriously high scores. Our recognition of this necessity does not mean that we doubt the accuracy of the reported data. To secure it, however, with the present apparatus was heartbreaking as most of our data had to be discarded or else we would have had to forfeit the exactness essential in parapsychological research. (Of the data discarded for technical reasons, most of them tended to support the psi hypothesis.)

Although the writers are confident of the accuracy of the data, it is recognized that such mere self-confidence must be something less than fully convincing to others. Therefore, the present data are offered only as preliminary and suggestive. It was not our purpose to design an experiment to prove the existence of psi—a task which has already received so much attention—but rather to see if our attempt to get at a level of relatively unconscious participation might be a fruitful mode of approach. Given this purpose, our data are encouraging. The observed effect from eight subjects, while not very strong, is consistent. If the expectations generated by our logic and our preliminary findings hold in further testing, it may signify the opening of a path for breaking through the high "noise level" which usually desensitizes psi in laboratory investigations, and for achieving the repeatability of findings so earnestly craved by workers in parapsychology.

References

1. McElroy, W. A., and Brown, W. R. K. Electric shock for errors in ESP card tests. *J. Parapsychol.*, 1950, **14**, 257-66.
2. Pratt, J. G., and Woodruff, J. L. Size of stimulus symbols in ESP. *J. Parapsychol.*, 1939, **3**, 121-58.
3. Rhine, J. B. *Extrasensory Perception*. Boston: Bruce Humphries, 1934.
4. Rhine, J. B., et al. *Extrasensory Perception After Sixty Years*. New York, Henry Holt, 1940.
5. Richet, Charles. *Thirty Years of Psychical Research*. New York, Macmillan, 1923.
6. Stuart, Charles E. The effect of rate of movement in card matching tests of ESP. *J. Parapsychol.*, 1938, **2**, 171-83.
7. Tyrrell, G. N. M. Further research in ESP. *Proc. Soc. psych. Res.* (London), 1937, **44**, 99-168.
8. Webster, D. D. An automatic testing and recording device for experiments in ESP. *J. Parapsychol.*, 1949, **13**, 107-17.

Department of Psychology
Brandeis University
Waltham, Massachusetts

CONTINUOUS VARIABLE TRIALS[1]

By R. A. McConnell

ABSTRACT: The results of some psi tests, such as ESP data obtained in tests with clock cards, may be evaluated from two points of view. On the one hand, the responses may be regarded as forming a continuous variable, i.e., the success of each trial is represented by the degree to which the call deviates from the target. On the other hand, the responses may be treated as discontinuous, each call being scored as either a hit on the target or a miss.

This paper examines the mathematics involved in assessing results on the basis of continuous and binomial (discontinuous) recording of results and compares the efficiency of the two methods under different circumstances. It is shown that in an experiment scorable by either method, the sensitivity of the method in detecting psi will depend upon how the psi process operates. If the hits are clustered close to the target, binomial scoring will be more sensitive. If they are widely spread, discontinuous scoring will be more efficient.

An appendix provides a table of the continuous-variable distribution function that will facilitate the use of this method of scoring.—Ed.

Most of the past tests for psi have used discontinuous variables (cards, dice) in which one of a number of mutually exclusive events can occur with a fixed probability. An experiment with such variables generally consists of a number of independent, identical trials (card guesses or die throws) and the simplest mathematical evaluation of the result is based upon binomial theory.

Numerous experiments have been proposed, and some have been carried out, in which the dependent variable is continuous rather than discontinuous in nature. For such situations it is usually possible to use either continuous or discontinuous analysis of the results; and the question naturally arises as to which method is more efficient for the detection of psi. For example, the pointer position of Knowles' psi experiment (3) could have been analyzed as a continuous variable. And in the clock-card experiments[2] of Fisk, Mitchell, and West (2, 6) the weighted discontinuous treatment of

[1] Biophysics Department publication No. 57. Presented in part at the National Biophysics Conference, Columbus, Ohio, March 4-6, 1957. This research has been financed by the A. W. Mellon Educational and Charitable Trust. The computational assistance of R. J. Snowdon in the preparation of the Appendix is gratefully acknowledged.

[2] Summarized in reference (4).

near hits suggests that continuous-variable analysis may be a sensitive method for detecting ESP.

We shall consider this problem in the following terms. Let the range of the continuous variable x representing the subject's response be 2, extending from -1 to $+1$ with the target point T at the origin, as shown in Fig. 1a. In any one trial the subject will respond at a point R with an absolute response error |R-T|, or simply $|x|$. The subject's task is to make |R-T| as small as possible. The chance probability of response is arranged to be constant throughout the response range.

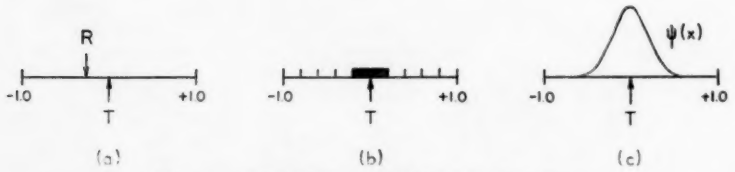


FIG. 1. Continuous and Binomial Scoring.

For example, in an experiment where the subject is asked to guess where a hidden, turning pointer comes to rest, the point T of Fig. 1a would correspond to the rest position of the pointer, and the range of x from -1 to $+1$ would correspond to angular positions of from $-180°$ to $+180°$ measured from that rest position. (The clock-cards, which allow only 12 discrete positions for target and response, can be considered as a discontinuous approximation to the continuous variable, and the conclusions from our discussion should apply *mutatis mutandis.*)

In the situation represented by Fig. 1a, two methods of scoring suggest themselves. In the first, a region of tolerance may be assumed on either side of T, a response falling within this region to be considered a target hit, and one falling outside, a miss. The chance distribution of the resulting trial scores will be binomial. In the second method the absolute response errors |R-T| for the several trials may be added to give the subject's score: the smaller the score the better the performance. This is the continuous-variable treatment, whose advantages relative to the binomial will be examined in this paper.

As a specific binomial example we shall consider the response range to be divided into ten equal parts, the two parts adjacent to T being considered hit regions, as indicated in Fig. 1b. The chance probability of a hit p is then $\frac{1}{5}$, the standard deviation of a trial $\sqrt{p(1-p)}$ is $\frac{2}{5}$, and the scoring of a run of 25 trials to test the usual null hypothesis will correspond exactly to the scoring of an "open" ESP card deck.

To compute the distribution parameters for continuous-variable scoring one integrates over the range of possible absolute response error. The chance-determined mean and standard deviation for a single trial are given by the formulas (1):

$$\mu = \int |x|\, \varphi|x|\, dx \qquad \sigma^2 = \int (|x| - \mu)^2\, \varphi|x|\, dx$$

where for the chance-determined situation the probability density $\varphi|x|$ equals 1 if the range of $|x|$ is from 0 to 1.

For n trials (binomial or continuous):

$$\mu_n = n\mu \qquad \sigma_n^2 = n\sigma^2 \qquad \sigma_n/\mu_n = \sigma/(\sqrt{n}\mu)$$

The mean and standard deviation are presented in Table 1 for several values of n and p, along with the coefficient of variation for a single trial. Notice that the coefficient of variation (σ/μ) of a binomial trial for $p = \frac{1}{5}$ is $\sqrt{12}$ times as large as that of a continuous-variable trial. To reach a given coefficient of variation in the total score requires 12 times as many binomial trials ($p = \frac{1}{5}$) as continuous-variable trials. Off hand, one might erroneously conclude that in situations scorable either way, continuous-variable scoring is 12 times as effective for the detection of psi as the customary binomial scoring. Actually, the relative merits of the two scoring methods can be assessed only by taking into account their relative sensitivity to psi in the manner to be described below.

In Appendix A the continuous-variable distribution function is tabulated and compared to the normal error curve for 5, 10, and 25 trials. This appendix will be useful to those who wish to score continuous-variable experiments.

Upon the assumption that some trials, constituting a fraction k of the total number of trials n, are "psi determined," a figure-of-merit may be calculated for comparing the efficiencies of continuous-variable and binomial scoring. A "psi-determined trial" is here defined as one in which $|x|$ follows a specified frequency distribution

Table 1

COMPARISON OF PARAMETERS OF CHANCE-CONTROLLED TRIALS FOR CONTINUOUS-VARIABLE AND BINOMIAL SCORING

Method of Scoring	Single Trial			25 Trials		
	μ	σ	σ/μ	Score range	μ	σ
Continuous						
Max. score/trial=½.....	0.25	0.14	0.58	0-12.5	6.25	0.72
Max. score/trial=1.....	0.50	0.29	0.58	0-25	12.50	1.44
Binomial						
$p=1/25$.....	0.04	0.20	4.90	0-25	1.00	0.98
$p=1/5$.....	0.20	0.40	2.00	0-25	5.00	2.00
$p=1/2$.....	0.50	0.50	1.00	0-25	12.50	2.50

determined by the operation of psi. (It will be assumed for simplicity that psi-determined trials, like chance-determined trials, are independent of one another.) The relative efficiency of the two scoring methods depends upon the nature of the distribution assumed for the psi-determined trials. However, this aspect of the problem will be considered after the figure-of-merit has been explained.

The figure-of-merit may be most easily understood in terms of the total score of n trials and the trial standard deviation ($\sqrt{p(1-p)}$ for the chance binomial case, and $1/\sqrt{12}$ for a chance-determined, continuous variable with an absolute range of 1.0). In either case it is possible to compute the mean absolute deviation $\triangle$ from the chance-expected score caused by the psi-determined trial. For binomial scoring *in the particular situation* where all psi-determined trials are hits, $\triangle$ is the psi-determined score minus the chance-expected score $(1-p)$. Of course, in the general situation a psi-determined trial might show a hit, a near miss, intentional avoidance, etc. For the continuous variable the deviation $\triangle$ per psi trial depends upon the assumed distribution of the psi-determined responses. And if the continuous-scoring method is to be comparable to the binomial, the latter must be allowed to depend upon that same distribution, as will be explained hereafter.

In any case, the total deviation from chance expectation in n trials is $nk\triangle$ and the standard deviation (under the assumption that psi is not operating) is $\sqrt{n}\sigma$. If the number of binomial trials is

sufficiently large so that both binomial and continuous scores can be evaluated on the normal curve, the critical ratio corresponding to a given criterion of significance will be for both methods $k\sqrt{n}\triangle/\sigma$. It is evident that $\triangle/\sigma$ is a suitable figure-of-merit for comparing the two methods of scoring since it will be inversely proportional to the square root of the total number of trials n to reach the chosen level of significance.

The fraction k, representing the number of psi-determined trials divided by the total number of trials, will presumably depend upon psychological factors, and these will vary with the nature of the experiment. Likewise, the frequency distribution which is followed by the psi-determined trials will depend upon psychological factors. Hence we must view with reserve any attempt to use the mathematical methods of this paper to determine the relative merit of a binomial experiment and a *different* continuous-variable experiment. However, in the particular case where the same experiment can be scored either way and the method of scoring is unknown to the subject, we may perhaps assume that both k and the distribution of psi-determined trials is independent of the scoring method. It is under these assumptions that we use $\triangle/\sigma$ as a figure-of-merit. We shall now evaluate $\triangle/\sigma$ for various binomial parameters and several psi-determined frequency distributions.

The probability density (i.e., frequency distribution) of the response is a constant for chance-determined trials as noted above, but for psi-determined trials it might have any conceivable shape. The most "reasonable" of these might be a normal distribution centered on the target in the manner of Fig. 1c. This would be the situation, for example, when the aim of the subject's ESP response was disturbed by a large number of small, independent, additive causes.

Since we are concerned in this paper only with the absolute magnitude of the response error, and not with its sign, we shall consider the sum of the left and right halves as a single distribution extending from T to + 1.0; and, as we have already done for the chance-determined case, we shall compute the mean and standard deviation of the *absolute* response error $|x|$, since these are what appear in the figure-of-merit. No generality in the conclusions is lost by re-

stricting our attention to a continuous-variable model with a fixed range of 1.

For continuous-variable scoring where ψ is the frequency distribution of the psi-determined trials:

$$\Delta = \left| \int_0^1 |x|\,\psi|x|\,dx - 0.5 \right|$$

$$\sigma = 1/\sqrt{12} \quad \text{(from Table 1)}$$

For binomial scoring:

$$\Delta = \left| \int_0^1 a(|x|)\psi|x|dx - p \right|$$

where the assigned score, $a(|x|) = \begin{cases} 1, & 0<|x|<p \\ 0, & |x|>p \end{cases}$

$$\sigma = \sqrt{p(1-p)}$$

For distributions of ψ whose mode is at $|x| = 0$, the general behavior of these integrals can be understood by assuming ψ to be the Gaussian curve of Fig. 1c. In Table 2, continuous and binomial scoring are compared for variable standard deviation of a normal ψ; and in Table 3, the two scoring methods are compared for variable binomial trial probability p.

In addition to figures-of-merit the "trial ratio" is tabulated. This is the ratio of the expected number of trials to reach a given level of significance when using the alternate scoring methods. For example, a trial ratio of 0.75 means that on the average only ¾ as many trials would be needed for binomial scoring as for continuous scoring to reach a chosen probability value.

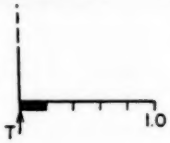
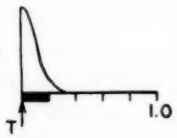
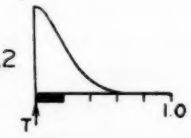
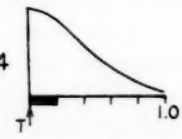
Discussion

It is evident from Tables 2 and 3 that for narrow distributions of psi-determined trials, binomial scoring is more efficacious than continuous; and the smaller p, the greater the advantage of the binomial method.

Within an increasingly wide distribution of psi-determined trials and fixed binomial p, the trial ratio (Table 2) goes through a minimum. This results from the fact that for continuous scoring, *any*

Table 2

FIGURES-OF-MERIT FOR CONTINUOUS AND BINOMIAL SCORING ($p = 1/5$)

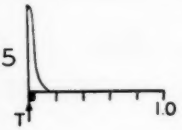
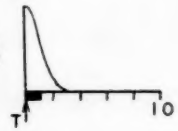
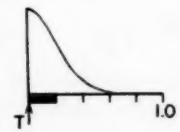
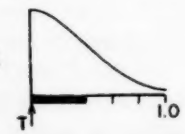
DISTRIBUTION OF PSI-DETERMINED TRIALS $\psi\ \|x\|$	CONTINUOUS SCORING Δ/σ	BINOMIAL SCORING Δ/σ	TRIAL RATIO $\left(\frac{\Delta}{\sigma}\right)^2_c \Big/ \left(\frac{\Delta}{\sigma}\right)^2_b$
$\sigma_\psi = 0$ 	1.73	2.00	0.75
$\sigma_\psi = 0.1$ 	1.46	1.89	0.60
$\sigma_\psi = 0.2$ 	1.18	1.21	0.95
$\sigma_\psi = 0.4$ 	0.63	0.46	1.88

response error is disadvantageous; while for binomial scoring, response errors smaller than p all receive full credit as a hit. The maximum advantage for binomial scoring will occur when ψ is a discontinuous line distribution ($\sigma_\psi = 0$) located with an $|x|$ infinitesimally less than p, so as to give the poorest possible continuous score for which all trials are binomial target hits. For this distribution ($p = 1/5$) the continuous-scoring figure-of-merit drops to 1.04 and the binomial remains at 2.00, giving a trial ratio of 0.27.

The distributions of psi-determined hits analyzed in Tables 2 and 3 are centered on the target. It is an obvious converse implication of the foregoing paragraph that a displacement of psi-determined responses away from the target point will give a decided advantage in sensitivity for continuous-variable scoring relative to binomial scoring when the displacement exceeds the size of the binomial hit

Table 3

FIGURES-OF-MERIT FOR CONTINUOUS AND BINOMIAL SCORING ($\sigma_\psi = p$)

DISTRIBUTION OF PSI-DETERMINED TRIALS $\psi \mid x \mid$	CONTINUOUS SCORING Δ/σ	BINOMIAL SCORING Δ/σ	TRIAL RATIO $\left(\frac{\Delta}{\sigma}\right)^2_c \Big/ \left(\frac{\Delta}{\sigma}\right)^2_b$
p = 1/25 	1.62	3.28	0.24
p = 1/10 	1.46	1.94	0.56
p = 1/5 	1.18	1.21	0.95
p = 2/5 	0.63	0.58	1.18

region p. Nevertheless, it does not seem proper to claim this as an intrinsic advantage for continuous-variable scoring; for if the displacement is of a consistent sort, the nominal target region can be reassigned and binomial scoring will thereby regain what advantage it may have. Or if the displacement is erratic, goodness-of-fit tests are appropriate, and in general more sensitive, than any method which specifies the target.

That the relative advantage in the scoring methods is generally insensitive to the exact form of the ψ distribution (providing the mode is at $|x| = 0$) can be seen from a distribution in which $\psi = 2.5$ for $0 \leqq |x| \leqq 0.4$ (a rectangular distribution). For $p = 1/5$ the figures-of-merit and trial ratio are 1.04, 0.75, 1.92. These figures are roughly comparable to those for a normal distribution with $\sigma_\psi = 0.3$.

In choosing between continuous and binomial scoring, a technical consideration arises in certain applications where the binomial-scoring method can introduce a bias that will not be compensated by Yates' correction for continuity. This might occur, for example, when measuring the response range and response error on a graphic tape record by means of a ruler, where the least count of the ruler is non-integrally related to the size of the "hit" region in a variable but non-random fashion. The correction for this bias would be uncertain and laborious.[3]

With the continuous variable, on the other hand, no bias will occur providing obvious precautions are observed (e.g., to conceal the target when reading the response, and vice versa). There will, of course, be a slight increase in the observed chance standard deviation over the theoretical chance standard deviation.[4]

In defining figure-of-merit a particular model was used. To what extent does the model represent the known nature of psi and to what extent are the conclusions drawn dependent upon the model? The answer to the first question is unclear, but that to the second provides the confidence we seek. The model is more general than at first glance appears.

A principal assumption was that psi operates for a given fraction k of the trials and in those trials response follows a determinate distribution. Since the trial ratio is independent of k, psi and non-psi responses may be lumped into one distribution without affecting the conclusions drawn, i.e., for present purposes the concept of k is unnecessary, even though it is illuminating and has been used by others in non-mathematical speculations about psi.

Hence the question which this paper answers is this: Given a hypothetical distribution of observed responses having certain

[3] For example, when applying a binomial probability of ⅙ to an absolute response interval which in size is exactly 12 of the smallest assignable scale units, one must count as hits all response errors measured as 1 and 9/10 of those measured as 2, under the assumption that the reader will read the response error $|x|$ to the nearest assignable scale unit. Such a correction is laborious at best, and impractical if the absolute response region varies in size.

[4] Following the example of the preceding footnote, and assuming that 50 per cent of the time the reader will correctly choose the nearest of the smallest assignable scale divisions, and that the remainder of the time he will be off by one division, the increase in the trial standard deviation caused by the measuring procedure will be less than 3 per cent.

plausible, characteristic departures from the chance distribution, which method of analysis is more sensitive for a null test, the continuous-variable or binomial analysis? This is strictly a mathematical question not involving the nature of psi.

Conclusions

In an experiment scorable by either the continuous-variable or binomial methods, the relative sensitivity to psi will depend upon the frequency distribution of the psi-determined responses. If the responses are clustered close to the target, binomial scoring will be more sensitive. If they are widely spread, the continuous-variable method should be used.

In an experiment in which the deviations are continuous in nature, continuous-variable scoring avoids a least-count bias which may arise when converting continuous data into binomial data.

Appendix A

The distribution function Φ for the sum of H of n identical variables with a constant probability density for values between 0 and 1 was found by Laplace (**5**, p. 277) to be of the form

$$\Phi(H) = \sum_{j} (-1)^j (H-j)^n / [j!\,(n-j)!]$$

where j assumes all integral values equal to and greater than zero, for which $(H-j)$ is positive.

This function has been computed for $n = 5$, 10, 25 and compared with the normal function in Table A. Unlike the binomial this distribution is unskewed. Evidently for most purposes one can use the normal error curve without correction in testing the usual null hypothesis with as few as 10 trials.

References

1. Arley, N., and Buch, K. R. Introduction to the *Theory of Probability and Statistics*. New York: John Wiley, 1950.
2. Fisk, G. W., and Mitchell, A. M. J. ESP experiments with clock cards: A new technique with differential scoring. *J. Soc. psych. Res.* (London), 1953, **37**, 1-14.
3. Knowles, E. A. G. On the susceptibility of a manually operated random selector to psi dexterity: A preliminary report. *J. Parapsychol.*, 1952, **16**, 41-58.

Table A

THE CONTINUOUS-VARIABLE DISTRIBUTION FUNCTION AND ITS RATIO TO THE NORMAL

5 Trials (σ=0.645498)			10 Trials (σ=0.912871)			25 Trials (σ=1.443378)		
Score	Φ_c	Φ_c/Φ_n	Score	Φ_c	Φ_c/Φ_n	Score	Φ_c	Φ_c/Φ_n
.00	0.0000	(5.4×10^{-5})	.0	0.0000	(2.2×10^{-8})	.0	0.0000	($2.4 \times 10^{-}$)
.25	8.1380×10^{-6}	0.033	.5	2.6912×10^{-10}	6.5×10^{-4}	1.0	6.4470×10^{-25}	8.0×10^{-11}
.50	2.6042×10^{-4}	0.268	1.0	2.7557×10^{-7}	0.047	2.0	2.1632×10^{-13}	1.2×10^{-6}
.75	1.9775×10^{-3}	0.590	1.5	1.5888×10^{-5}	0.252	3.0	5.4570×10^{-14}	2.3×10^{-3}
1.00	8.3333×10^{-3}	0.828	2.0	2.7943×10^{-4}	0.551	4.0	7.1221×10^{-11}	0.037
1.25	2.5391×10^{-2}	0.962	2.5	2.4692×10^{-3}	0.800	5.0	1.7415×10^{-8}	0.171
1.50	6.1979×10^{-2}	1.022	3.0	1.3463×10^{-2}	0.946	6.0	1.3742×10^{-6}	0.411
1.75	1.2689×10^{-1}	1.035	3.5	5.0452×10^{-2}	1.006	7.0	4.6234×10^{-5}	0.667
2.00	2.2500×10^{-1}	1.026	4.0	1.3890×10^{-1}	1.016	8.0	7.8073×10^{-4}	0.857
2.25	3.5347×10^{-1}	1.012	4.5	2.9452×10^{-1}	1.009	9.0	7.3537×10^{-3}	0.960
2.50	5.0000×10^{-1}	1.000	5.0	5.0000×10^{-1}	1.000	9.5	1.8575×10^{-2}	0.986
						10.0	4.163×10^{-2}	1.000
						10.5	8.338×10^{-2}	1.005
						11.0	1.5029×10^{-1}	1.006
						11.5	2.454×10^{-1}	1.005
						12.0	3.654×10^{-1}	1.002
						12.5	5.0000×10^{-1}	1.000

NOTES

Parenthetical values are Φ_n.

Values are correct to within 1 in the final place.

4. MITCHELL, A. M. J., AND FISK, G. W. Clock cards and differential scoring techniques. *J. Parapsychol.*, 1954, **18**, 153-64.
5. USPENSKY, J. V. *Introduction to Mathematical Probability.* New York: McGraw-Hill, 1937.
6. WEST, D. J., AND FISK, G. W. A dual ESP experiment with clock cards. *J. Soc. psych. Res.*, London, 1953, **37**, 185-97.

Department of Biophysics
University of Pittsburgh
Pittsburgh, Pennsylvania

BOOK REVIEWS

TELEPATHIE EN HELDERZIENDHEID (Telepathy and Clairvoyance). By W. H. C. Tenhaeff. Zeist: W. de Haan, Antwerp etc.: Standaard Boekhandel, 1958. 157 pp.

This treatise is one of the Phoenix Pocketbooks, forming part of a series in the Dutch language. As such, one can welcome it as another endeavor to bring parapsychology to the general public. The booklet is written in an entertaining way and makes a point of giving many concrete, illuminating examples. The thread running through the various chapters is the notion of *memory*. To begin with, all of us are capable of remembering things (in German: sich erinnern); then there is the enhanced recollection in hypnosis; sometimes there may be a "remembering" of what other persons have stored in their memory (in German one might call this ein sich *innern*). Precognition amounts to a "remembering" (innern) of the future; the seeming contact with the dead might be a remembering of the contents of *their* memory. Reincarnation is the purported recollection of past lives. Here the author deals with the cases of Bridey Murphy, Shanti Devi, and others.

One of the critical remarks we are inclined to make is that Professor Tenhaeff, who does not accept reincarnation, devotes more pages (31) to this problem than to an exposition of the quantitative method in parapsychology (15 pages). Nevertheless, he describes the latter as an important pioneer contribution, whereas he warns against over-rating it (p. 152). Elsewhere (p. 58) he writes that the late Professor Heymans assumed a definite connection between cryptomnesia [hidden memory] and *fausse reconnaissance* [mistaken recognition]. We know, however, of no passage where Heymans does so; he only connected *fausse reconnaissance* with depersonalization (*Gesammelte Kleinere Schriften III*, 1927, p. 1 ff.; *Zeitschrift für Psychologie* 36, 1904, p. 321 ff.). Tenhaeff also mentions depersonalization in this context, but the step from depersonalization to cryptomnesia is his, not Heymans', and the reason for this step is not very evident to us. Our final remark is:

Is it quite correct to use exactly the same title for this book as was used for the book published in 1936 by the other Dutch parapsychologist, Dr. P. A. Dietz, the existence of which is well known to Professor Tenhaeff (p. 144)?

J. J. POORTMAN

PSI COGNITION. By K. Ramakrishna Rao (with a foreword by J. B. Rhine) Tenali, India: Tagore Pub. House, 1957. 179 pp.

Psi Cognition is a revised version of a Master's thesis written by Mr. Rao for the Philosophy Department at Andhra University in Waltair, India, where he is Lecturer in Philosophy. The book is intended as a brief introduction to the evidence for ESP and some of the theories which have been put forth to explain it. Of special interest is the discussion in Chapter II of Indian theories of Yogic perception.

Since the book is based on a philosophical thesis, the main emphasis is on theoretical questions rather than on experimental work, and the survey of actual research is inclined to be sketchy and imprecise.

Mr. Rao's discussion of Yogic systems of perception constitutes the principal new material in his book and is quite worth separate mention. The author indicates that most schools argue that by sheer contemplation one can know everything, while one school, the Mimamsa, teaches that one should be concerned with knowing nothing. But as far as I understand the matter, the kind of perception developed by intense meditation is not really specific perception of objects or ideas, but rather a general undifferentiated intuition of life. If, as Patanjali and Bhoja say, all things in all times can be known through Yoga, why hasn't someone who has learned all these things told us what some of them are? The answer to this question is rather obvious, but it also expresses the reason why mystical intuition is unlikely to have anything to do with ESP. Mr. Rao is properly circumspect about claiming any direct connection between Yoga and ESP. He is content to indicate that the two areas do not contradict one another.

Mr. Rao has made a notable effort to put parapsychology before Indian readers. In discussing a great deal of relevant material in

so few pages he is sure to whet the appetites of many who will want to consult the various papers and books to which he refers. His book is a healthy sign of recent interest in parapsychology in India. It is unfortunate that he has been put to a disadvantage by a wholly disproportionate number of printers' errors and other inaccuracies. I feel that he would have done a better job if he had stuck a little closer to ground level throughout the book, but these considerations should not detract too much from the general value of his contribution nor from our hopes that he will continue to help parapsychology gain a foothold in a country where so many fertile minds are not so steeped in crass materialism as ours.

STEPHEN I. ABRAMS

PARAPSYCHOLOGICAL ABSTRACTS

SARA R. FEATHER

[This section consists of abstracts of reports published in other journals and of unpublished reports filed at the Parapsychology Laboratory at Duke University. Readers may obtain unabridged copies (Thermo-Fax) of the unpublished reports by writing to the *Journal of Parapsychology*, College Station, Durham, N. C. The price is ten cents per page.]

BERG, F., v.d., JONG, A. H., de, and KAPPERS, J. Voortgezet onderzoek naar het bestaan van objectieve helderzienheid met het toestel van Mak. (A further investigation of the existence of pure clairvoyance with the apparatus of Mr. Mak.) *Tijdschrift voor Parapsychologie*, 1957, **25**, 224-32.

An apparatus for testing clairvoyance was constructed in such a way that the percipient could choose between a white or a black marble which would drop into the appropriate tray when he pressed one of two handles. The marbles in the apparatus were completely invisible and issued at random. One hundred and thirty-three percipients did 40,000 guesses from which there were 19,949 hits (CR =0.51). Later on, the experiments were continued with a second group of percipients whose scoring was higher (CR>1.9). No significant results were obtained. It was observed that the men were inclined to press the right-hand handle more often than the left-hand one.—*G. Zorab*

CRUMBAUGH, J. C. Are negative ESP results attributable to traits and attitudes of subjects and experimenters? (Unpublished manuscript.)

As a test of the hypothesis that traits and attitudes of both subject and experimenter can influence the results in ESP tests, a study was factorially designed to manipulate the following personality and attitudinal variables in both subjects and experimenters: self-confidence, insecurity, belief in ESP, and disbelief in ESP. Sixteen subjects and 16 experimenters were selected from the

author's psychology classes on the basis of their marked deviation from the norms on the I (self-confidence) scale of the Guilford-Martin GAMIN and on a questionnaire measuring attitude toward psi phenomena. Marginally significant differences were obtained, generally in the direction predicted. However, a second experiment designed to repeat the basic features of this experiment, with some modification, gave results which were at the chance level of expectation, as were the results of both experiments combined. The author considers this outcome to be "an offense to the ESP hypothesis" and maintains that the validity of the ESP hypothesis must remain in question until an experiment can be devised which "consistently yields results favorable to ESP in most of the repetitions done by qualified skeptical experimenters who undertake it."

NASH, C. B., AND NASH, C. S. Checking success and the relationship of personality traits to ESP. *J. Amer. Soc. psych. Res.*, 1958, **52**, 98-107.

Thirty-four sophomore biology students from St. Joseph's College scored on the chance level in a total of 289 runs directed at recorded target digits during individual testing sessions. Significant results (P=.013) were obtained, however, in 1714 additional trials made within the run to identify the calls that were correct. Correlation analysis revealed that the success in checking hits was not the result of a higher frequency of checks when the run score was high, although both high-scoring and low-scoring subjects checked a higher percentage of hits than misses. The manner in which the targets were recorded had been varied, unknown to the subjects, in five different ways, but the results appeared to be unaffected by this fact. No significant relationship appeared between the scoring of the 24 subjects who completed the 10-run experiment and their degree of belief in ESP as measured on a five-point questionnaire administered before and after the experiment. Their hits were also examined as to correlation with: (1) time used to complete the run, (2) a measure of self-confidence, and (3) scores on 10 personality traits of the Guilford-Zimmerman Temperament Survey. None of these correlations was independently

significant. Although the comparison was not planned in advance of the experiment, a study of the relation of ESP scores to the available percentile ranks for the American Council on Education Psychological Examination gave a correlation coefficient of +0.55 ($P<.01$).

OTANI, S. Studies on the influence of the mental and physiological conditions upon ESP function. (Unpublished manuscript.)

Some suggestion that a change in skin resistance may be related to ESP scoring was indicated in a clairvoyance experiment conducted at Chiba University, Japan. The author and O. Shioya acted alternately as subject and experimenter. Otani had a significant score in runs in which he exhibited a change in psychogalvanic skin resistance of from 10,000 to 15,000 ohms during the run. When the change in skin resistance was above or below this range, his scores were in a negative direction. Shioya's scores did not show these differences. Contrary to earlier published findings, Otani's scoring showed no relationship to whether the mental set was one of conscious concentration or one of relaxed indifference.

WEISS, D. L. A college student's research in parapsychology. (Unpublished manuscript.)

A study of the relationship between GESP and extraversion-introversion is reported, along with the results of other exploratory tests the author conducted for her own interest while a freshman in college. This study involved grouping 30 subjects into the 10 most introverted, the 10 most extraverted, and a middle group of 10, on the basis of their ratings on the Freyd and Heidbetter introversion-extraversion self-report questionnaire. The introverted group obtained a deviation of −11 from chance in 60 GESP runs (six runs per subject), whereas the extraverted group obtained a deviation of +51, and the middle group a deviation of +7. The difference between the introverted group and the extraverted group is significant (CR of the difference=2.82). (Although this difference is not in the direction predicted by the author, it is consistent with the findings of previous ESP and introversion-extraversion studies.) The report includes a number of observations into the psychological conditions affecting success in GESP.

THE PARAPSYCHOLOGICAL ASSOCIATION

Proceedings of the First Convention

The first convention of the Parapsychological Association was held in Townsend Harris Hall, the City College of New York, September 4-6, 1958. The program included seventeen papers, an invited address by Dr. Gardner Murphy, a business meeting, two Council meetings, a banquet, and a Presidential Address by Dr. Robert A. McConnell.

The Program Committee consisted of Miss Margaret Anderson, Chairman, Dr. J. G. Pratt, and Dr. Carroll B. Nash. The members of the Arrangements Committee were Mr. Robert Taetzsch, Chairman, and Dr. Karlis Osis.

Program

Thursday Morning, September 4

9:00 Registration
10:00 Convention Call to Order...Dr. R. A. McConnell, President
Secretary's Report.....Miss Rhea White, Secretary
Treasurer's Report.....Mr. W. G. Roll, Treasurer
11:00 Council Meeting

Thursday Afternoon, September 4
Session Chairman: Dr. Michael Scriven

1:30 "Emotional Telepathy" by Dr. George H. Estabrooks (Department of Psychology, Colgate University)

Emotion would seem to be an important factor in many spontaneous cases of psychic phenomena. The following exploratory experiment was aimed at the objective measurement of this factor. The experimenters chose pairs of subjects on the basis of a presumed sympathy or rapport between them: identical twins, fiancés, mother-son, and others. One subject acted as "sender," the other as "receiver." They were situated in rooms 50 feet apart, and to each subject was attached a psychogalvanic reflex instrument. The indicating recorders were on one panel in the room of the sender. Mild

stimuli were administered to the sender—a pinch, a hair-pull, a sudden noise. The experimenters sought for concomitant variations on the two instruments, but results to date are negative. This is to be expected because of the mild nature of the stimuli. The experimenters will now proceed to strong emotional states under hypnosis; also to the use of drugs.

2:00 "Physiological Conditioning for Psi Performance" by Dr. J. H. Rush (4021 19th St., Boulder, Colorado) and Mr. Harold A. Cahn (University of Colorado)[1]

The Boulder, Colorado, psi research group is seeking ways to develop psi performance in persons who do not normally manifest it. Experiments with mescaline in three subjects indicate both psi facilitation and relationships between the subject's personality type and subjective effects of the drug. Perceptual channels that are most used in conscious processes seem least susceptible to confusion and hallucination. This observation suggests that a verbal-minded person might best try ESP through visualization. Access to the unconscious was vividly demonstrated in one subject who re-experienced an impression he had received unconsciously. Subjects failed in card-guessing tests but showed encouraging success in tests with free material, particularly token objects. These findings are exploratory and tentative. The experimenters expect to investigate other methods of inducing a detached state and to extend the use of Rorschach and other personality tests. They will also investigate correlations between EEG records and subjective states.

2:30 "The Unconsciousness of ESP" by Mr. W. G. Roll (Parapsychology Laboratory, Duke University)

If ESP occurrences are to be described in terms of a perception model, we must assume that the objects which are used as ESP targets have causal properties other than those which produce sound waves, light waves, etc. However, responses to these properties—unlike responses to visual stimuli, say—are not associated with a characteristic conscious experience. If there is an introspective experience, it comes in terms of one of the sense modalities. On the other hand, the experiences associated with ESP and sensory

[1] Presented by Mr. W. S. Cox.

perception resemble each other in that neither by itself can be distinguished from experiences which have a wholly subjective origin, such as dreams and hallucinations. When a sensory or an extrasensory experience is identified for what it is, this seems to be because of the subject's ability to infer the origin of his experience by rational means.

3:30 "Terminal Salience with Multiple Digit Targets" by Dr. Carroll B. Nash (Parapsychology Laboratory, St. Joseph's College, Philadelphia)

Fifty-five subjects took part in a clairvoyance experiment designed to test the hypothesis that a multiple-aspect target is extrasensorially perceived as a unit. The targets were numbers containing one, two, three, four, or five randomly selected digits. There were no significant differences between the scores with the five kinds of target. The results did show a terminal salience effect in that the scores on the outer digits of the three-, four-, and five-digit numbers were higher than those on the inner digits.

4:00 "Paracausal Systems and Synchronicity" by Stephen I. Abrams (University of Chicago)

C. G. Jung's studies on synchronicity represent the first attempt to integrate the findings of parapsychology into a general theory of personality. Synchronicity is a concept denoting acausal but meaningful relations linking events on the basis of some presumed harmony between aspects of reality. Within the context of Jungian psychology, strict repeatability of psi experiments appears unlikely because synchronistic manifestations seem to be neither consciously observable nor controllable, but follow an apparently random order of their own. These considerations have far-reaching implications for the research worker, but one would be premature in attempting to draw any conclusions from them at this time.

Thursday Evening, September 4

7:00 Convention Dinner Address: "Progress in Parapsychology" by Dr. Gardner Murphy (Menninger Foundation, Topeka, Kansas)[2]

[2] This address is given in full elsewhere in this number of the *Journal*.

Friday Morning, September 5

Session Chairman: Mr. Wadih Saleh

9:00 "An Investigation of ESP Between Teachers and Pupils in First and Second Grades" by Mr. J. G. van Busschbach (Kleverlaan 151, Haarlem, Holland)[3]

Earlier schoolroom tests in Holland used fifth- and sixth-grade pupils as subjects. The present report deals with a continuation of these tests in 40 classes of first- and second-grade pupils in Amsterdam (1956) and Dordrecht (1957). In the experiment the teacher, concealed and located in the rear of the classroom, looked at a card selected at random from a set of five, while the pupils attempted to identify the card. Three different types of target symbols were used: figures, objects, and colors. The significant results which were obtained were primarily due to the scoring of the first-grade girls.

9:30 "GESP Test with First-Grade Pupils" by Mrs. Christiane Vasse (136 Boulevard de Chateaudun, Amiens, France)[4]

A series of clairvoyance tests were carried out with first-grade pupils in a French school to find out whether the teacher-pupil relationship developed during the first year in school is conducive to a high ESP performance. Three series of tests were conducted—one near the end of the 1954-55 school year; one at the start of the next school year; and a third at the end of that year. The two series conducted near the end of the school year gave highly significant positive results. The results were negative in the series coming at the start of the term.

10:00 "The Relationship Between Changes in Student Attitude and ESP Scoring" by Miss Rhea White (Parapsychology Laboratory, Duke University)[5]

This report covers the fifth of a series of experiments by Anderson and White to test the relationship between teacher-pupil attitudes and the level of scoring in clairvoyance tests. The special point of

[3] Presented by Miss Margaret Anderson.

[4] Based on a report by Christiane and Paul Vasse in the *Journal of Parapsychology*, Sept., 1958.

[5] Based on a report by Margaret Anderson and Rhea White in the *Journal of Parapsychology*, Sept., 1958.

this experiment was to see if changes in teacher-pupil attitudes from the beginning of the school semester to the end in a particular Senior classroom would be reflected in the ESP scores. The results show that changes in the teacher's attitude toward the pupil, as well as changes in their mutual attitudes, were not associated with significant changes in ESP scores. However, changes in the attitudes of the students toward the teacher were significantly related to changes in the level of ESP scoring. The findings also indicate that one might expect better results in ESP tests administered toward the end rather that at the beginning of the school term.

11:00 "GESP Tests with Fourth- and Fifth-Grade Pupils" by Mr. James A. Hall (Southwestern Medical School, Dallas)[6]

A series of four GESP tests were given to all fourth- and fifth-grade students of the Gladewater, Texas, public schools to determine if there is any relationship between teacher-pupil attitudes and the pupils' ESP scores under conditions permitting either telepathy or clairvoyance. The pupils' regular classroom teacher, hidden by a screen, acted as agent. Target symbols were either colors, vowels, or arithmetic symbols. The teachers were asked to rank students into three groups: those whom they would choose, exclude, or have no particular feeling about including in their "ideal" class. Pupils were asked to evaluate their teacher on the basis of a preference test modified from that published by Margaret Anderson. The ESP data, separated according to these criteria, did not give extrachance results.

11:30 "The Relationship Between ESP Scoring and Student Class Grade" by Miss Margaret Anderson (Parapsychology Laboratory, Duke University)

A series of 10 ESP studies in schools have been carried out with the pupils as subjects. One objective was to see if a relationship could be observed between students' class grades and their level of ESP scoring. Two of the studies were tests for precognition; eight were for clairvoyance. The data concern 39 teachers of 44 different classes in nine different grade levels. It was found that students with high scholastic grades, A and B, scored above mean

[6] Presented by Miss Rhea White.

chance expectation in their ESP tests, while students with low grades, D and E, had scores that were below. The difference between the ESP scores of the two groups is statistically significant.

Friday Afternoon, September 5

Session Chairman: Dr. Jan Ehrenwald

1:30 "ESP and the Bird Navigation Problem" by Dr. J. G. Pratt (Parapsychology Laboratory, Duke University)

Extensive research on long-distance orientation in birds has been conducted in several centers during the past eight years. Not only has the existence of a true navigational ability independent of directional training been established, but the results have contributed to knowledge of this behavior in a number of respects. However, no solution of the problem in terms of a specific sensory hypothesis has been found. This behavior therefore remains an area of interest and challenge for the parapsychologist, although fruitful research will depend on the development of an adequate experimental method.

2:00 "A Test of Adding Extrasensorially Perceived Digits" by Mrs. Catherine S. Nash (Parapsychology Laboratory, St. Joseph's College, Philadelphia)

An experiment was performed to determine whether subjects can add numbers that are knowable only by ESP. There were 10 different targets: the digits from 1 to 5, and five problems in adding two digits whose sum was equal to a digit from 1 to 5 (such as 0 + 1). There were significant results on the single-digit targets, but only chance results on the targets requiring addition. The results indicate that for this experiment, at least, ESP and the addition of numbers was not a diametric or unitary function.

2:30 "An Exploratory Study of Some Personality Correlates Associated with PK Performance" by Dr. Robert L. Van de Castle (Department of Psychology, Idaho State College)[7]

An exploratory study was made to discover whether some of the personality correlates related to ESP scoring might also be

[7] Presented by Dr. Gertrude R. Schmeidler.

related to PK scoring. Thirty-one college students were administered a battery of personality tests, and each subject completed 18 PK runs. No significant over-all deviation or decline effects were found for the PK results as a whole. However, when analyzed according to personality classifications, several encouraging differences in scoring were found. A suggestive positive correlation occurred between expansive-compressive ratings and PK scores. Seven of the eight subjects rated as "completely spontaneous" by Rorschach measures scored above chance and seven of the eight "completely inhibited" subjects, below chance. The study also includes the Rosenzweig P-F test, an intelligence test, and sheep-goat ratings.

3:00 "Exploratory 'Scaled' PK Placement Tests with Nine College Students, With and Without Distance" by Mr. Jarl Fahler (Parapsychology Foundation)

A series of dice tests were carried out with nine college students as subjects, and using two different test situations. In the first, the subjects were in the PK room where they were able to see the mechanical dice-throwing apparatus and follow the whole procedure. In the other situation, they were in another room, more than 20 yards away, and were unable to see the table or the moving cubes. In both situations, the subject himself released the dice by pressing a button at the end of an electric cord. Attempts were made, with statistically significant results, to predict whether the subjects would score positively or negatively on the basis of earlier tests of their ESP abilities.

4:00 Presidential Address: "Scaled Measurement in Psi Research" by Dr. R. A. McConnell (University of Pittsburgh)[8]

Saturday Morning, September 6

Session Chairman: Dr. C. J. Ducasse

9:00 "Three New ESP Test Machines and Preliminary Results" by Mr. W. C. Stewart (College of Engineering, Duke University)

[8] This address is given in full elsewhere in this number of the *Journal*.

Three Senior students in electrical engineering invented electronic devices for testing ESP. This report gives the results of an exploratory ESP experiment in which these machines were used. A contest was held on the Duke University campus with cash prizes for the highest scores. There were 99 contestants, each of whom had 50 trials on each machine. On one of the machines (the author's) the over-all scores were above chance average to a significant degree; on one they were at chance; and on the third, below.

9:30 "Application of Statistical Quality Control Techniques to Statistical Psi Control Problems," by Mr. Robert Taetzsch (Bakelite Company, Bound Brook, N. J.)

Practical applications of psi may be possible at this time, even considering the present low yields, if control techniques are developed to a high degree. The problems of statistical psi control are analogous to problems in statistical quality control dealt with in industry. Examples are given of how techniques used in the latter may be applied to statistical psi control problems.

10:00 "Some Avoidable Misconceptions of Parapsychology" by Dr. J. B. Rhine (Parapsychology Laboratory, Duke University)[9]

Three misconceptions of what has been established by the investigations of psi are considered. The first is the opinion that only "empty correlations" have resulted. The second is the idea that ESP is claimed to be an enduring capacity of the individual and species. Third, and most pervasive, is the notion that psi has been represented as a phenomenon demonstrable or repeatable under specifiable conditions. Serious errors result from these misconceptions. One of the consequences of the third is the demand that a showdown demonstration be produced before the case for psi be taken seriously by science; a second, the claim that valid negative conclusions can be drawn from experiments which give no evidence of psi; and a third, that all experimental results on a given problem must be statistically combined to avoid unwitting selection. How-

[9] Presented by Dr. J. G. Pratt.

ever, a correct formulation of what has actually been established would invalidate all these views and help to avoid the consequent mistakes in practice.

11:00 Business Meeting

11:30 Council Meeting

GLOSSARY

A description of the basic experimental methods, of the findings, and of the statistical procedures for evaluating ESP and PK results may be found in *Parapsychology, Frontier Science of the Mind* by J. B. Rhine and J. G. Pratt (published by Charles C. Thomas, Springfield, Illinois, U.S.A., Blackwell Scientific Publications, Ltd., Oxford, England, and The Ryerson Press, Toronto, Canada).

AGENT: The "sender" in tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

BT (Basic Technique): The clairvoyance technique in which each card is laid aside by the experimenter as it is called by the subject. The check-up is made at the end of the run.

CALL: The ESP symbol selected by the subject in trying to identify a target.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

Chance Expectation = Mean Chance Expectation: The most likely score if only chance variation is present.

Chance Average: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic cognition of the mental activities of another person.

CR (Critical Ratio): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearsons.)

CR_d (Critical Ratio of the Difference): The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DISPLACEMENT: ESP responses to targets other than those for which the calls were intended.

Backward Displacement: ESP responses to targets preceding the intended targets. Displacement to the targets one, two, three, etc. places preceding the intended target are designated as (—1), (—2), (—3), etc.

Forward Displacement: ESP responses to targets coming later than the intended targets. Displacement to the targets one, two, three, etc. places after the intended target are designated as (+1), (+2), (+3), etc.

DT (Down Through): The clairvoyance technique in which the cards are called down through the pack before any are removed or checked.

ESP (Extrasensory Perception): Awareness of or response to an external event or influence not apprehended by sensory means.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, square, cross and waves (three parallel wavy lines).

ESP PACK: Twenty-five ESP cards.

Closed Pack: An ESP pack composed of five each of the five symbols.

Open Pack: An ESP pack made up of the ESP symbols selected in random order.

EXPECTATION: See ***Chance.***

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

GESP (General Extrasensory Perception): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

MCE (Mean Chance Expectation): see *Chance.*

P (Probability): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with behavioral or personal effects that are demonstrably nonphysical (that is, which do not fall within the scope of known physical principles).

PERCIPIENT = SUBJECT: The person who makes the calls in an ESP test.

PK (Psychokinesis): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PLACEMENT TEST: A PK technique in which the aim of the subject is to try to influence falling objects to come to rest in a designated part of the throwing surface.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring responses to free material. A judge ranks the stimulus objects (usually pictures in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes which are nonphysical in nature. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI-MISSING: Exercise of psi ability in a way that avoids the target the subject is attempting to hit.

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

PSYCHICAL RESEARCH: Older term used for parapsychology.

QD (Quarter Distribution): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RANDOM ORDER: An order of events which displays no trends or regularities that would allow any inference regarding one event from one or more of the others in the series.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in any given unit of trials.
Total Score: Total of scores made in a given number of runs.
Average Score: Total score divided by number of runs.

SD (Standard Deviation): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$ in which n is the number of single trials, p the probability of success per trial, and q the probability of failure.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. The criterion commonly used in parapsychology today is a probability value of .01 or less, or a deviation in either direction such that the CR is 2.58 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

SPONTANEOUS PSI EXPERIENCE: Natural, unplanned occurrence of an event or experience that seems to involve parapsychical ability.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

STM (Screened Touch Matching): An ESP card test in which the subject indicates in each trial (by pointing to one of five key positions) what he thinks the top card is in the inverted pack held by the experimenter behind a screen. The card is then laid opposite that position.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP (including precognition) the objective or mental events the subject is attempting to respond to. In PK tests, the faces of the die which the subject attempts to bring up in the act of throwing. In placement tests, the designated part of the throwing surface.

Target Card: The card which the percipient is attempting to identify or otherwise indicate a knowledge of.

Target Pack: The pack of cards the order of which the subject is attempting to identify.

Target Face: The face on the die which the subject tries to turn up by PK.

TELEPATHY: Extrasensory perception of the mental activities of another person (as distinguished from clairvoyant cognition of objective events).

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.

HERE AND THERE IN THE LITERATURE

Parapsychology is a new field, and the problem is to decide whether it's new like phrenology or new like physiological psychology, whether it has a great future as an historical example or as a respectable subject. Add to its novelty the fact that it is by definition the repository of the hopes of behavioral anti-scientists, religious scientists, and superstitious nonscientists, and one begins to understand the usual reactions of psychologists. On the whole, however, it is about time they began to understand their *own* reactions and stopped acting like stock examples for a theory of socially conditioned perception. The height-of-amused-tolerance school favors the argument that there are still *lots* of things to be done in the field of normal perception before we have time to waste looking into the science fiction area. The let's-not-rush-into-this-thing group feels we should really wait until something more substantial appears. The absolutely-open-minded-but-haven't-you-heard bunch trots out a story about how the statistics has been disproved in *Nature* or about a lab assistant from Duke who has told all. And so on. Well, colleagues, if there's anything to this stuff, it is more important to psychology than relativity was to physics; if there's nothing to it, the statistical underpinning of ordinary psychology is overdue for collapse; and, whatever there is *to* it, there's an awful lot *of* it, and it won't come out and talk to you, so you will actually have to go read some of it before you can seize on those little anecdotes as salvation.

—Michael Scriven, "On Even Keel
Through Parapsychology"
Contemporary Psychology, October, 1958